

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 November 2002 (21.11.2002)

PCT

(10) International Publication Number
WO 02/092626 A2

- (51) International Patent Classification⁷: C07K 14/34 (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (21) International Application Number: PCT/EP02/05359
- (22) International Filing Date: 15 May 2002 (15.05.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
01111774.4 15 May 2001 (15.05.2001) EP
01130253.6 20 December 2001 (20.12.2001) EP
- (71) Applicant (*for all designated States except US*): CELLZOME AG [DE/DE]; Meyerhofstrasse 1, 69117 Heidelberg (DE).
- (72) Inventor; and
- (75) Inventor/Applicant (*for US only*): NEUBAUER, Gitta [DE/DE]; Cellzome AG, Meyerhofstrasse 1, 69117 Heidelberg (DE).
- (74) Agent: GRÜNECKER, KINKELDEY, STOCKMAIR & SCHWANHÄUSSER; Maximilianstrasse 58, 80538 München (DE).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 02/092626 A2

BEST AVAILABLE COPY

(54) Title: CLEAVAGE AND POLYADENYLATION COMPLEX OF PRECURSOR MRNA

(57) Abstract: The present invention relates to novel components of the cleavage/polyadenylation machinery of precursor mRNA as well as to the complex containing the new components and its use. The complex is obtained by using one component thereof as a bait and isolating a highly organised complex consisting of at least 13 distinct proteins.

CLEAVAGE AND POLYADENYLATION COMPLEX OF PRECURSOR MRNA

2. BACKGROUND OF THE INVENTION

Polyadenylation of precursor mRNA (pre-mRNAs) is an obligatory step in the maturation of most eukaryotic transcripts. The addition of poly(A) (polyadenosine) tails promote transcription termination and export of the mRNA from the nucleus. Furthermore, the poly(A) tails have the function to increase the efficiency of translation initiation and to help to stabilize mRNAs. Polyadenylation occurs posttranscriptionally in the nucleus of eukaryotic cells in two tightly coupled steps: the endonucleolytic cleavage of the precursor and the addition of a poly(A) tail.

In the yeast *Saccharomyces cerevisiae*, the pre-mRNA 3'-end processing signals are not as well conserved as in mammalian cells (see below). In addition to the cleavage and polyadenylation site, two cis-acting elements, called the efficiency element and the positioning element, are found upstream of the cleavage site. Efficiency elements contain the sequence UAUUAU (or close variants thereof) and are often repeated. The sequence AAUAAA and several related sequences can function as a positioning element.

Fractionation of yeast extracts led to the separation of protein factors that are required for mRNA 3'-end formation in vitro. The cleavage reaction requires cleavage factors I and II (CF I and CF II), whereas polyadenylation involves CF I, polyadenylation factor I (PF I) and poly(A) polymerase (Pap1).

CF I can be separated into two activities, CF IA and CF IB. CF IA is needed for both processing steps and is a heterotetrameric protein with subunits of 38, 50, 70 and 76 kDa that are encoded by the RNA5, CLP1, PCF11 and RNA14 genes. Rna14 shares significant sequence similarity to the 77 kDa subunit of mammalian cleavage stimulation factor (CstF) and Rna15 contains a RNA-binding domain homologous to that of the 64 kDa subunit of CstF.

In addition to the above mentioned four CFI subunits, Pab1 (poly(A) binding protein) was identified in purified CFI fractions. Both biochemical and genetic data indicate an involvement of Pab1 in poly(A) length control. CF IB consists of a single protein called Nab4/Hrp1 and is required for cleavage site selection and polyadenylation.

A multiprotein complex which has CFII-PF I (= CPF) activity consists of nine polypeptides: Pap1 (poly(A) polymerase), Pta1, Pfs1, Pfs2, Fip1, Cft1/Yhh1, Cft2/Ydh1, Ysh1/Brr5, and Yth1. Pap1, a 64 kDa protein, was the first component of the yeast 3'-end formation complex to be purified to homogeneity. Pta1 is a 90 kDa protein which is required for both cleavage and polyadenylation of mRNA precursors. Pfs2 is a 53 kDa protein that contains seven WD40 repeats. Pfs2 has been shown to directly interact with subunits of CFII-PF1 and CFIA and is thought to function in the assembly and stabilization of the 3'-end processing complex. Fip1 has been demonstrated to physically interact with Pap1, Yth1 and Rna14 and it is believed that it tethers Pap1 to its substrate during polyadenylation. Cft1/Yhh1, Cft2/Ydh1, Ysh1/Brr5, and Yth1 are the counterparts of the four subunits of the mammalian cleavage and polyadenylation specificity factor, CPSF160, CPSF100, CPSF73 and CPSF30, respectively.

Furthermore TIF4632 has been found to interact with Pab1 (see Table 1)

For the mammalian system, various data have been presented which have given evidence both for a conserved mechanism and also showed some differences between the yeast and the mammalian structures.

The composition and function of the mammalian complex based on the data to date is as follows:

The cleavage and polyadenylation factor (CPSF) is composed of 4 subunits: CPSF160 (involved in mRNA and poly(A) polymerase (PAP) binding), CPSF100, CPSF 73 and CPSF30 (involved in mRNA and PABII binding).

CPSF binds the AAUAAA hexanucleotides. CPSF links the mRNA 3'-end processing to the transcription. CPSF exists as a stable complex with the transcription factor TFIID complex. The 160 kDa subunit of CPSF binds to several hTAFII. TFIID recruits CPSF to

the RNA polymerase II pre-initiation complex. Upon transcriptional activation CPSF dissociates from TFI_{II} and associates with the elongating RNA pol II (CTD carboxy-terminal domain of the largest subunit of the RNA polymerase II). CPSF is thought to travel with RNA pol II until they reach the polyadenylation site, where CPSF can bind the AAUAAA element. CPSF is required for the termination of transcription.

The interaction between CPSF and the AAUAAA element is weak and not so specific. The binding of CPSF to the hexanucleotide is greatly enhanced by a 2nd component of the poly-adenylation machinery, the cleavage stimulation factor (CstF), which binds the G-U rich motif. CstF also binds the RNA pol II through its 50 kDa subunit (CstF50). Furthermore, CstF50 binds another component of the transcriptional machinery: BRCA1 associated RING domain protein (BARD1). BARD1 also interacts with RNA pol II. BARD1-CstF50 interaction inhibits polyadenylation in vitro and may prevent inappropriate mRNA processing during transcription. CstF is composed of 3 subunits: CstF64 (binds mRNA and symplekin (yeast homolog: Pta1), CstF77 (binds CPSF160, CstF64, CstF50) and CstF50 (binds RNA pol II and BARD1). The co-operative binding of CPSF and CstF to the polyadenylation site forms a ternary complex, which functions to recruit the other components of the polyadenylation machinery to the cleavage site: the two cleavage factors (CFI_m and CFI_{II}) and the poly(A) polymerase (PAP).

CFI_m is an heterodimer of 4 subunits 72, 68, 59, 25 components: one essential, CFI_mA and one stimulatory, CFI_mB. CFI_mA contains hPCF11p and hClp1p (binds cPSF and CFI_I). CFI_mB contains no factors previously shown to be involved in 3'-end processing and may be a new 3'-end processing factor. Although the identity of the proteins that perform the cleavage step is still unknown, it is well established that both CFI_m and CFI_{II} are required. The reaction products of the cleavage suggest that a metal ion is involved. Surprisingly, PAP (but not its catalytic activity) is required for the cleavage.

After the cleavage step CstF, CFI_m and CFI_{II} are dispensable. PAP bound to CPSF (through its 160 kD subunit) can start polyadenylating the cleaved 3'-end, but at that step, the process is very inefficient. The poly(A) binding protein II (PAB II) can bind the nascent poly(A) chain as soon as it reaches a minimal length of 10 poly(A). PAB II also interacts with the CPSF30. The binding of PAB II greatly stabilizes PAP at the 3'-end of the mRNA, supporting the progressive synthesis of a long poly(A) tail. In the nucleus,

the length of the poly(A) tail is restricted to about 250 poly(A). This size restriction is probably achieved through stoichiometric binding of multiple PAB II. It is not yet known how the incorporation of a certain amount of PAB II in the complex terminates processive elongation.

CstF is part of the mammalian 3'-end processing complex and is a heterotrimeric protein with subunits of 77, 64 and 50 kDa. CstF-50 has been shown to interact with the BRCA1-associated protein BARD1 and this interaction suppresses the nuclear mRNA polyadenylation machinery in vivo. In a recent study it was found that treatment of cells with DNA damage-inducing agents causes a transient, but specific, inhibition of mRNA 3'-end processing in cell extracts. This inhibition reflects the BARD1/CstF interaction and involves enhanced formation of a CstF/BARD1/BRCA1 complex. A tumor-associated germline mutation in BARD1 decreases binding to CstF-50 and renders the protein inactive in polyadenylation inhibition. These results support the existence of a link between mRNA 3'-end formation and DNA repair/tumor suppression. The in vivo function of these interactions may be to inhibit the cleavage and polyadenylation of pre-mRNAs on polymerase molecules that are stalled at sites of DNA repair.

Cleavage stimulation factor (CstF) is one of the multiple factors required for mRNA polyadenylation in mammals. CstF-64 may play a role in regulating gene expression and cell growth in B cells. The concentration of one CstF subunit (CstF-64) increases during activation of B cells, and this is sufficient to switch IgM heavy chain mRNA expression from membrane-bound to secreted form. Reduction in CstF-64 causes reversible cell cycle arrest in G0/G1 phase, while depletion results in apoptotic cell death.

In contrast to what is observed in yeast, the sequence elements in mammals, which specify the site of cleavage and polyadenylation, flank the site of endonucleolytic attack. One is the hexanucleotide AAUAAA found 10-30 bases upstream of the cleavage/polyadenylation site. The second is a G-U-rich motif located 20-40 bases downstream of the cleavage/polyadenylation site. These two elements and their spacing determine the site of cleavage/polyadenylation and also the strength of the polyadenylation signal.

Some other elements, like sequences upstream of the AAUAAA (upstream sequence elements, USEs) play regulatory roles.

A schematic presentation of the motifs underlying mammalian polyadenylation and yeast polyadenylation are shown in Fig. 1. A review on the formation of mRNA 3'-ends in eukaryotes is given in Zhao, Hyman and Moore in *Microbiology and Molecular Biology Reviews*, 1999, pp. 405-445. A comparison of mammalian and yeast pre-mRNA 3'-end processing is also given in Keller and Minvielle-Sebastia in *Nucleus and gene expression in Current Opinion in Cell Biology*, 1997, Vol. 9, pp. 329-336.

There are diseases which involve defects in the function of the polyadenylation machinery.

Many viruses interact directly with components of the mRNA processing machinery. The herpes simplex virus type 1 (HSV-1) immediate early (alpha) protein ICP27 is an essential regulatory protein that is involved in the shutoff of host protein synthesis. It affects mRNA processing at the level of both polyadenylation and splicing. During polyadenylation, ICP27 appears to stimulate 3' mRNA processing at selected poly(A) sites. The opposite effect occurs on host cell splicing. That is, during HSV-1 infection, an inhibition in host cell splicing requires ICP27 expression. This contributes to the shutoff of host protein synthesis by decreasing levels of spliced cellular mRNAs available for translation. A redistribution of splicing factors regulated by ICP27 has also been seen.

Epstein-Barr virus BMLF1 gene product EB2 seems to affect mRNA nuclear export of intronless mRNAs and pre-mRNA 3' processing. EB2 contains an Arg-X-Pro tripeptide repeated eight times, similar to that described as an mRNA-binding domain in the herpes simplex virus type 1 protein US11.

Interestingly, both viruses have been found to precede the onset of lymphomas.

Influenza A virus NS1A protein binds the 30 kDa subunit of the cleavage and polyadenylation specificity factor (CPSF), NS1 protein (NS1A protein) via its effector domain targets the poly(A)-binding protein II (PABII) of the cellular 3'-end processing machinery. In vitro the NS1A protein binds the PABII protein, and in vivo causes PABII

protein molecules to relocate from nuclear speckles to a uniform distribution throughout the nucleoplasm. In vitro the NS1A protein inhibits the ability of PABII to stimulate the processive synthesis of long poly(A) tails catalyzed by poly(A) polymerase (PAP). Such inhibition also occurs in vivo in influenza virus-infected cells. Consequently, although the NS1A protein also binds the 30 kDa subunit of the cleavage and polyadenylation specificity factor (CPSF), 3' cleavage of some cellular pre-mRNAs still occurs in virus-infected cells, followed by the PAP-catalyzed addition of short poly(A) tails. Subsequent elongation of these short poly(A) tails is blocked because the NS1A protein inhibits PABII function. The NS1 effector domain functionally interacts with the cellular 30 kDa subunit of CPSF, an essential component of the 3' end processing machinery of cellular pre-mRNAs.

Metachromatic leukodystrophy (MLD) is a lysosomal storage disorder caused by the deficiency of arylsulfatase A (ASA). A substantial ASA deficiency has also been described in clinically healthy persons, a condition for which the term pseudodeficiency was introduced. The mutations characteristic for the pseudodeficiency (PD) allele have been identified. Sequence analysis revealed two A-G transitions. One of them changes the first polyadenylation signal downstream of the stop codon from AATAAC to AGTAAC. This causes a severe deficiency of a 2.1-kilobase (kb) mRNA species. The deficiency of the 2.1-kb RNA species provides an explanation for the diminished synthesis of ASA seen in pseudodeficiency fibroblasts.

MLD patients have been identified who are homozygous for the ASA-PD allele and it is thought that the allele might play a role in the development and progression of disease.

There is a tight link between cell cycle control and polyadenylation machinery suggesting an important role of this machinery in the development of cancer. Cyclin-dependent enzymes seem to regulate the activity of the polyadenylation machinery. The amounts of some factors of the mRNA 3' processing machinery (CstF) increase in mitotically active cells in phases of the cell cycle preceding DNA synthesis. The amount of the 64-kDa subunit CstF-64 increases 5-fold during the G0 to S phase transition and concomitant proliferation induced by serum in 3T6 fibroblasts. The increase in CstF-64 is associated with the G0 to S phase transition. Cdc2-cyclin B phosphorylates PAP at the Ser-Thr-rich region.

However, as it seems now, most diseases associated with defects in mRNA processing are caused by mutations in cis-acting elements that disrupt sequences essential for pre-mRNA splicing. These can be canonical sequences at the intron-exon border or located within an exon. They directly affect the expression of a single mutated gene. Approximately 15% of the nucleotide substitutions that cause human diseases disrupt pre-mRNA splicing. Thus these diseases do not seem to be directly caused by alterations in the polyadenylation/cleavage-machinery.

However, since recently evidence for a number of interrelationships between polyadenylation/cleavage and splicing is accumulating (for review see Zhao, Hyman and Moore in Microbiology and Molecular Biology Reviews, 1999, pp. 405-445), it might very well be that alterations in the 3'-end processing machinery contribute to the etiology of these diseases.

Examples of diseases caused by incorrect splicing are mentioned below:

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease involving degeneration of cortical motor neurons and spinal/bulbar motor neurons. In the sporadic form of the disease, the neuron degeneration is caused by excessive extracellular glutamate. The glutamate transporter functional in the CNS is the astrocyte EAAT2 which is altered in ALS. The pre-mRNA for EAAT2 is aberrantly spliced in the brain regions affected. The reason for this is still unknown, but the defect lies probably in one or a few auxiliary splicing factors that regulate the splicing of a sub-set of pre-mRNA in these cells. The factors have not yet been identified.

The human papillomavirus (HPV) E2 protein plays an important role in transcriptional regulation of viral genes as well as in viral DNA replication. HPV-5 (an EV epidermodysplasia verruciformis-HPV) protein can specifically interact with cellular splicing factors including a set of prototypical SR proteins and two snRNP-associated proteins (Lai, Teh et al. 1999, J. Biol. Chem. 274, pages 11832-41). Interestingly all these three viruses have been associated with cancer progression. Papillomavirus infection precedes cervical cancer, whereas EBV and HSV-8 have been described in association with lymphomas.

In hepatocellular carcinoma, there is a defect in mRNA splicing. In this disease, there are anti-nuclear antibodies to a 64 kD protein, which has splicing factor motifs. A defect in the regulatory subunit 3 of the protein phosphatase 1(PP1) has been found in haematological malignancies and in lung, ovarian, colorectal and gastric cancers. Low PP1 activity has been observed also in acute myelogenous leukaemia.

Heterogeneous nuclear ribonucleoproteins (hnRNP) associate with pre-mRNA and have a role in RNA processing and splice site selection. HnRNP A2 shows a marked overexpression in lung cancer and brain tumours and has thus been used as a biomarker for these tumor types.

The development of antinuclear antibodies (ANA) in malignancies has been described but its mechanism is still not understood. A great diversity of ANA specificities is found in hepatocellular carcinoma. In hepatoma sera antibodies co-localize with non-snRNP splicing factor SC35, suggesting that the antigenic targets might be involved in mRNA splicing. Hepatocellular carcinoma has a significantly higher frequency of ANA than chronic hepatitis C, chronic hepatitis B, alcoholic liver cirrhosis or healthy donors.

In some autoimmune diseases, a possible link has been detected to a preceding virus infection, like Epstein-Barr virus in SLE. Furthermore it seems that even vaccination is potentially dangerous: a candidate for cytomegalovirus CMV vaccine is glycoprotein gB (UL55). Immunization with an adenovirus-gB construct (Ad-gB) not only induces a significant anti-viral response, but a significant IgG auto-antibody response ($p > 0.005$) to the U1-70 kDa spliceosome protein. Auto-antibodies to U1-70 kDa are part of the anti-ribonucleoprotein response seen in systemic lupus erythematosus and mixed connective tissue disease.

At least two molecules which are also part of the complex are known to be inhibited by natural toxins or treatment against various diseases.

Protein phosphatase 1 is inhibited by several natural product toxins.

The marine toxins include the cyanobacteria-derived cyclic heptapeptide microcystin-LR and the polyether fatty acid okadaic acid from dinoflagellate sources. They bind to a common site on PP1. The dephosphorylation of PP1 is inhibited (among other serine/threonine phosphatases PP2A, PP2B, PP2C and PP5/T/K/H) by Fumonisin B1 (FB1), a mycotoxin produced by the fungus *Fusarium moniliforme*. This is a common contaminant of corn, and is suspected to be a cause of human esophageal cancer. FB1 is hepatotoxic and hepatocarcinogenic in rats, although the mechanisms involved have not been clarified.

Viral proteins are able to interfere with PP1 activity:

The transcription factor EBNA2 of the Epstein-Barr virus induces the expression of LMP1 onco-gene in human B- cells. EBNA2A from an EBV-immortalized B-cell line co-immunopurifies with a PP1-like protein. A PP1-like activity in nuclear extracts from EBV-immortalized B-cell line can be inhibited by a GST-EBNA2A fusion product.

Poly(A)polymerase (PAP) is affected by anticancer drugs and is inhibited by some antiviral agents.

Anticancer drugs:

Most anticancer drugs act through the mechanism of apoptosis. Apoptosis may be regulated at all levels of gene expression including the addition of the poly(A) tail to the 3' end of mRNAs. Drug combinations are more effective than single drugs and various chemotherapeutic strategies have therefore been developed. Dimethylsulfoxide (DMSO) in combination with interferon (IFN) results in pronounced PAP dephosphorylation, activity reduction and apoptosis of HeLa cells.

Purine and pyrimidine analogues often affect PAP activity. They are potentially useful agents for chemotherapy of cancer diseases. The anticancer drugs 5-Fluorouracil (5-FU), interferon and tamoxifen mediate both partial dephosphorylation and inactivation of poly(A) polymerase (PAP).

PAP (from isolated hepatic nuclei) is inhibited by cordycepin 5'-triphosphate. The nucleoside analogue cordycepin is a therapeutic agent for TdT+ (terminal

deoxynucleotidyl transferase positive) leukemia. In the presence of an adenosine deaminase inhibitor, deoxycoformycin (dCF), cordycepin is cytotoxic to leukemic TdT+ cells. A cordycepin analog of (2'-5') oligo(A) which can be synthesized enzymatically from cordycepin 5'-triphosphate and the core cordycepin analog can replace human fibroblast interferon in preventing the transformation of human lymphocytes after infection with Epstein-Barr virus B95-8 (EBV). The core cordycepin analog is not cytotoxic to uninfected lymphocytes and proliferating lymphoblasts.

Not only is PAP affected by anticancer drugs, but it has a possible use as a tumor marker involved in cell commitment and/or induction of apoptosis and could be used to evaluate tumor cell sensitivity to anticancer treatment.

Antiviral drugs:

Ara-ATP (arabinofuranosyladenosine triphosphate) is an antiherpetic drug that inhibits herpes simplex virus replication. It inhibits poly(A) polymerase activity by competing with ATP. It blocks both cleavage and polyadenylation reactions by interacting with the ATP-binding site on poly(A) polymerase, the activity of which is essential for the cleavage reaction.

Purine and pyrimidine analogues are also used as antiviral agents. As an example, the most extensively used drug against HSV is idoxuridine, the 5'-amino analog of thymidine.

A decrease in herpes simplex virus transcription and perturbation of RNA polyadenylation is induced by 5'-amino-5'-deoxythymidine (AdThd).

The cleavage stimulation factor (CSTF):

Treatment with hydroxyurea or ultraviolet light strongly, but transiently, inhibits 3' cleavage. This is accompanied by increased amounts of a CstF/BARD1/BRCA1 complex, though the amount of these proteins remains the same.

Despite the large body of information already available from the prior art concerning the cleavage/polyadenylation machinery of precursor mRNA up to now not all components of the machinery are known not to speak of the composition of the complex as a whole.

3. SUMMARY OF THE INVENTION

An object of the present invention was to identify the components of the cleavage/polyadenylation machinery of precursor mRNA and provide new components of the cleavage/polyadenylation machinery to provide the machinery and to provide new targets for therapy.

By applying the process according to the invention to the isolation of the polyadenylation/cleavage machinery from yeast 32 new components could be identified which are Act1 (SEQ ID:1), Cka1 (SEQ ID:7), Eft2 (SEQ ID 11), Eno2 (SEQ ID: 13), Glc7 (SEQ ID:15), Gpm1 (SEQ ID:17), Hhf2 (SEQ ID:21), Hta1 (SEQ ID:23), Hsc82 (SEQ ID:25), Imd2 (SEQ ID:27), Imd4 (SEQ ID:29), Met6 (SEQ ID:31), Pdc1 (SEQ ID:39), Pfk1 (SEQ ID:41), Ref2 (SEQ ID:47), Sec13 (SEQ ID:53), Sec31 (SEQ ID:55), Ssa3 (SEQ ID:57), Ssu72 (SEQ ID: 59), Taf60 (SEQ ID:61), Tkl1 (SEQ ID:65), Tsa1 (SEQ ID: 67), Tye7 (SEQ ID: 69), Vid24 (SEQ ID:71), Vps3 (SEQ ID: 73), Ycl046w (SEQ ID: 79), Ygr156w (SEQ ID: 81), Yhl035c (SEQ ID:83), Ykl018w (SEQ ID:85), Ylr221c (SEQ ID: 87), Yml030w (SEQ ID:91) and Yor179c (SEQ ID:93).

Said object is further achieved by the characterisation of Ycl046w (SEQ ID: 79), Ygr156w (SEQ ID: 81), Yhl035c (SEQ ID:83), Ykl018w (SEQ ID:85), Ylr221c (SEQ ID: 87), Yml030w (SEQ ID:91) and Yor179c (SEQ ID:93) as components of the cleavage/polyadenylation machinery.

The invention thus relates to:

An isolated complex selected from complex (I) and comprising

(a) a first protein, or a functionally active fragment or functionally active derivative thereof, which first protein is selected from the group of proteins in Table 1, column A, or a mammalian homolog thereof, or a variant of said protein encoded by a nucleic acid that hybridizes to the nucleic acid of said protein or its complement under low stringency conditions; and

(b) a second protein, or a functionally active fragment or functionally active derivative thereof, which second protein is selected from the group of proteins in Table 1, column B, or a mammalian homolog thereof, or a variant of said protein encoded by a nucleic acid that hybridizes to the nucleic acid of said protein or its complement under low stringency conditions, wherein said first protein and said second protein are members of a native cellular Polyadenylation-complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C and a complex (II) comprising at least two second proteins.

Furthermore, the invention relates to an isolated complex comprising all proteins in column C of table 1, or the mammalian homologs of those proteins, or variants of said proteins encoded by nucleic acid that hybridises to the nucleic acid of any of said proteins or its complements under low stringency conditions, wherein proteins are members of a native cellular complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

Furthermore, the invention relates to an isolated complex that comprises all but 1,2,3,4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17,18,19,20,21,22,23,24,25,26,27 or 28 of all proteins in column C of table 1.

Furthermore, the invention relates to the complex as described above comprising a functionally active derivative of said first protein and/or a functionaliy active derivative

comprising said first protein or said second protein fused to an amino acid sequence different from the first protein or second protein, respectively.

In a preferred embodiment of the present invention, the protein components of the complex are vertebrate homologs of the yeast proteins, or a mixture of yeast and vertebrate homolog proteins. In a more preferred embodiment, the protein components of the complex are mammalian homologs of the yeast proteins, or a mixture of yeast and mammalian homolog proteins. In particular aspects, the native component proteins, or derivatives or fragments of the complex are obtained from a mammal such as mouse, rat, pig, cow, dog, monkey, human, sheep or horse. In another preferred embodiment, the protein components of the complex are human homologs of the yeast proteins, or a mixture of yeast and human homolog proteins. In yet another preferred embodiment, the protein components of the complex are a mixture of yeast, vertebrate, mammalian and/or human proteins.

Furthermore, the invention relates to a complex as described above of claim that is involved in the 3' end processing activity. Such a complex might also exist as a module or subcomplex of a larger physiological protein complex or assembly.

Furthermore, the invention relates to a complex as described above comprising a fragment of said first protein and/or a fragment of said second protein, which fragment binds to another protein component of said complex.

Furthermore, the invention relates to a complex as described above, wherein the functionally active derivative is a fusion protein comprising said first protein or said second protein preferentially fused to an affinity tag or label.

It is further directed to complexes comprising a fusion protein which comprises a component of the complex or a fragment thereof linked via a covalent bond to an amino acid sequence different from said component protein, as well as nucleic acids encoding the protein, fusions and fragments thereof. For example, the non-component protein portion of the fusion protein, which can be added to the N-terminal, the C-terminal or inserted into the amino acid sequence of the complex component can comprise a few amino acids, which provide an epitope that is used as a target for affinity purification of the fusion protein and/or complex.

Furthermore the invention relates to a process for processing RNA comprising the step of bringing into contact any of the complexes described above with RNA, such that RNA is processed.

Furthermore, the invention relates to an antibody or a fragment of said antibody containing the binding domain thereof, which binds the complex as described above of claim and which does not bind the first protein when uncomplexed or the second protein when uncomplexed.

Furthermore, the invention relates to a pharmaceutical composition comprising the protein complex as described above and a pharmaceutically acceptable carrier.

Moreover, the present invention provides a process for the identification and/or preparation of an effector of a composition according to the invention which process comprises the steps of bringing into contact the composition of the invention or of a component thereof with a compound, a mixture of compounds or a library of compounds and determining whether the compounds or certain compounds of the mixture or library bind to the composition of the invention and/or a component thereof and/or affects the biological activity of such a composition or component and then optionally further purifying the compound positively tested as effector by such a process.

A major application of the composition according to the invention results in the identification of an active agent capable of binding thereto. Hence, the compositions of the invention are useful tools in screening for new pharmaceutical drugs.

Furthermore, the invention relates to a method for screening for a molecule that modulates directly or indirectly the function, activity, composition or formation of the complex as described above comprising the steps of :

- (a) exposing said complex, or a cell or organism containing said complex to one or more candidate molecules; and
- (b) determining the amount of, the 3' end processing activity for mRNA of, or protein components of, said complex, wherein a change in said amount, activity, or protein components relative to said amount, activity or protein components in the absence of said candidate molecules indicates that the molecules modulate function, activity or composition of said complex.

Furthermore, the invention relates to a method as described above, wherein the amount of said complex is determined.

Furthermore, the invention relates to a method as described above, wherein the activity of said complex is determined.

Furthermore, the invention relates to a method as described above, wherein said determining step comprises isolating from the cell or organism said complex to produce

said isolated complex and contacting said isolating complex with the substrate under conditions conducive to binding to the complex.

Furthermore, the invention relates to a method as described above, wherein the protein components of said complex are determined.

Furthermore, the invention relates to a method as described above, wherein said determining step comprises determining whether any of the proteins listed in column B of table 1 of said complex or the mammalian homologs thereof, or variant of said proteins encoded by a nucleic acid that hybridises to the nucleic acids of any of said proteins or its complements under low stringency conditions, is present in the complex, wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

Furthermore, the invention relates to a method as described above, wherein said method is a method of screening for a drug for treatment or prevention of diseases and disorders, preferably diseases or disorders such as infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis and cancer.

Furthermore, the invention relates to a method for screening for a molecule that binds the complex as described above comprising the following steps:

- (a) exposing said complex, or a cell or organism containing said complex, to one or more candidate molecules; and
- (b) determining whether said complex is bound by any of said candidate molecules.

Furthermore, the invention relates to a method for diagnosing or screening for the presence of a disease or disorder or a predisposition for developing a disease or disorder in a subject, which disease or disorder is characterized by an aberrant amount of, the 3' end processing activity for mRNA biochemical activity of, or component composition or formation of, the complex as described above, comprising determining the amount of, the 3' end processing activity for mRNA of, or protein components of, said complex in a sample derived from a subject, wherein a difference in said amount, activity, or protein components of, said complex in an analogous sample from a subject

not having the disease or disorder or predisposition indicates the presence in the subject of the disease or disorder or predisposition.

Furthermore, the invention relates to a method as described above, wherein the amount of said complex is determined.

Furthermore, the invention relates to a method as described above, wherein the activity of said complex is determined.

Furthermore, the invention relates to a method as described above, wherein said determining step comprises isolating from the cell or organism said complex to produce said isolated complex and contacting said isolated complex with the substrate under conditions conducive to binding to the complex.

Furthermore, the invention relates to a method as described above, wherein the protein components of said complex are determined.

Furthermore, the invention relates to a method as described above, wherein said determining step comprises determining whether any of the proteins listed in column B of table 1 of said complex or the mammalian homologs thereof, or variant of said proteins encoded by a nucleic acid that hybridises to the nucleic acids of any of said proteins or its complements under low stringency conditions, is present in the complex, wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

Furthermore, the invention relates to a method for treating or preventing a disease or disorder characterized by an aberrant amount of, the 3' end processing activity for mRNA of, or component composition or formation of, the complex as described above, comprising administering to a subject in need of such treatment or prevention a therapeutically effective amount of one or more molecules that modulate the amount of, the 3' end processing activity for mRNA of, or protein components or formation of, said complex.

Furthermore, the invention relates to a method as described above, wherein said disease or disorder involves decreased levels of the amount or activity of said complex. Furthermore, the invention relates to a method as described above, wherein said disease or disorder involves increased levels of the amount or activity of said complex.

Furthermore, the invention relates to the use of a molecule that modulates the amount of, the 3' end processing activity for mRNA of, or protein components or formation of the complex as described above for the manufacture of a medicament for the treatment or prevention of a disease or disorder, preferably diseases or disorders such as infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis; cancer

Furthermore, the invention relates to a kit comprising in one or more containers (a) an isolated first protein, or a functionally active fragment or functionally active derivative thereof selected from the proteins in column A of table 1 of a given complex or a mammalian homolog thereof, or a variant of said protein encoded by a nucleic acid that hybridises to the nucleic acid of said protein or its complement under low stringency conditions; and (b) an isolated second protein, or a functionally active fragment or functionally active derivative thereof selected from the proteins in column B of table 1 of a given complex or a mammalian homolog thereof, or a variant of said protein encoded by a nucleic acid that hybridises to the nucleic acid of said protein or its complement under low stringency conditions, wherein said first and said second protein are members of a native cellular complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

Furthermore, the invention relates to a kit comprising in a container the isolated complex as described above or the antibody as described above, optionally together with further reagents and working instructions. The further reagents may be, for example, buffers, substrates for enzymes but also carrier material such as beads, filters, microarrays and other solid carries. The working instructions may indicate how to use the ingredients of the kit in order to perform a desired assay..

Furthermore, the invention relates to such kits for use in processing of RNA and for use in the diagnosis, prognosis and screening in or for the diseases mentioned above.

Furthermore, the invention relates to a complex as described above, or the antibody or fragment as described above, for use in a method of diagnosing a disease or disorder, preferably the diseases or disorders as mentioned above.

Furthermore, the invention relates to a method for the production of a pharmaceutical composition comprising carrying out the method as described above to identify a molecule that modulates the function, activity or formation of said complex, and further comprising mixing the identified molecule with a pharmaceutically acceptable carrier.

Furthermore, the invention relates to a process for preparing complex as described above and optionally the components thereof comprising the following steps: expressing such a protein in a target cell, isolating the protein complex which is attached to the tagged protein, and optionally disassociating the protein complex and isolating the individual complex members.

Furthermore, the invention relates to the process as described above characterized in that the tagged protein comprises two different tags which allow two separate affinity purification steps.

Furthermore, the invention relates to the process as described above, characterized in that two tags are separated by a cleavage site for a protease.

Furthermore, the invention relates to a component of the said complex obtainable by a process as described above.

The present invention further relates to a composition, preferably a protein complex, which is obtainable by the method comprising the following steps: tagging a protein as defined above, i.e. a protein which forms part of a protein complex, with a moiety, preferably an amino acid sequence, that allows affinity purification of the tagged protein and expressing such protein in a target cell and isolating the protein complex which is attached to the tagged protein. The details of such purification are described in WO 00/09716 and in Rigaut, G. et al. (1999), Nature Biotechnology, Vol. 17 (10): 1030-1032 and further herein below. The tagging can essentially be performed with any moiety which is capable of providing a specific interaction with a further moiety, e.g. in the sense of a ligand receptor interaction, antigen antibody interaction or the like. The tagged protein can also be expressed in an amount in the target cell which comes close to the physiological concentration in order to avoid a complex formation merely due to high concentration of the expressed protein but not reflecting the natural occurring complex.

In a further preferred embodiment, the composition is obtained by using a tagged protein which comprises two different tags which allow two different affinity purification steps. This measure allows a higher degree of purification of the composition in question. In a further preferred embodiment the tagged protein comprises two tags that are separated by a cleavage site for a protease. This allows a step-by-step purification on affinity columns.

Furthermore, the invention relates to a complex as described above and/or protein thereof as a target for an active agent of a pharmaceutical, preferably a drug target in the treatment or prevention of disease or disorder, preferably diseases or disorders as mentioned above..

Furthermore, the invention relates to the proteins Ycl046w (SEQ ID: 59), Ygr156w (SEQ ID: 61), Yhl035c (SEQ ID:63), Ykl018w (SEQ ID:179), Ylr221c (SEQ ID: 67), Yml030w (SEQ ID:69), and Yor17c (SEQ ID:71), the mammalian homologs/orthologs of said proteins and functionally active fragments and derivatives of said proteins and the mammalian homologs thereof carrying one or more amino acid substitutions, deletions and/or additions and the nucleic acid encoding said proteins or said homologs, orthologs and functionally active fragments and derivatives thereof.

Such a nucleic acid may be used for example to express a desired tagged protein in a given cell for the isolation of a complex or component according to the invention. Such a nucleic acid may also be used for the identification and isolation of genes from other organisms by cross species hybridization.

The present invention further relates to a construct, preferably a vector construct, which comprises a nucleic acid as described above. Such constructs may comprise expression controlling elements such as promoters, enhancers and terminators in order to express the nucleic acids in a given host cell, preferably under conditions which resemble the physiological concentrations.

The present invention further relates to a host cell containing a construct as defined above.

Such a host cell can be, e.g., any eukaryotic cell such as yeast, plant or mammalian, whereas human cells are preferred. Such host cells may form the starting material for isolation of a complex according to the present invention.

Animal models and methods of screening for modulators (i.e., agonists, and antagonists) of the amount of, activity of, or protein component composition of, a complex of the present invention are also provided.

3.1 DEFINITIONS

The term "mammalian homolog" or "homologous gene products" as used herein means a component protein of the cleavage/polyadenylation machinery of a mammal which performs the same function as the corresponding yeast protein. Such homologs are also termed "orthologue gene products". The algorithm for the detection of orthologue gene pairs from yeast and mammalian and human uses the whole genome of these organisms. First, pairwise best hits are retrieved, using a full Smith-Waterman alignment of predicted proteins. To further improve reliability, these pairs are clustered with pairwise best hits involving *Drosophila melanogaster* and *C. elegans* proteins. Such analysis is given, e.g., in Nature, 2001, 409:860-921. The mammalian homologs of the yeast proteins according to the invention can either be isolated based on the sequence homology of the yeast genes to the mammalian genes by cloning the respective gene applying conventional technology and expressing the protein from such gene, or by isolating the mammalian proteins by isolating the analogous complex according to methods commonly known in the art, and as described in Section 6, *infra*.

The term "protein complex machinery" as used herein means a complex of proteins in the cell that is able to perform one or more functions of the wild type protein complex. The protein complex may or may not include and/or be associated with other molecules such as nucleic acid, such as RNA or DNA, or lipids.

As used herein, the term "percent identity" means the number of identical residues as defined by an optimal alignment using the Smith-Waterman algorithm divided by the length of the overlap multiplied by 100. The alignment is performed by the search program (W.R. Pearson, 1991, Genomics 11:635-650) with the constraint to align the maximum of both sequences.

As used herein, the term "derivatives" or "analogs of component proteins" or "variants" include, but are not limited, to molecules comprising regions that are substantially homologous to the component proteins, in various embodiments, by at least 30%, 40%, 50%, 60%, 70%, 80%, 90% or 95% identity over an amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to a sequence encoding the component protein under stringent,

moderately stringent, or nonstringent conditions. It means a protein which is the outcome of a modification of the naturally occurring protein, by amino acid substitutions, deletions and additions, respectively, which derivatives still exhibit the biological function of the naturally occurring protein although not necessarily to the same degree. The biological function of such proteins can e.g. be examined by available in vitro cleavage/polyadenylation assays as will be described below.

As used herein, the term "Therapeutics" includes, but are not limited to, a protein complex of the present invention, the individual component proteins, and analogs and derivatives (including fragments) of the foregoing (e.g., as described hereinabove); antibodies thereto (as described hereinabove); nucleic acids encoding the component protein, and analogs or derivatives, thereof (e.g., as described hereinabove); component protein antisense nucleic acids, and agents that modulate complex formation and/or activity (i.e., agonists and antagonists).

"Target for therapeutic drug" means that the respective protein (target) can bind the active ingredient of a pharmaceutical composition and thereby changes its biological activity in response to the drug binding.

"Effector of the cleavage/polyadenylation of precursor mRNA" means a compound that is capable of binding to a member of the cleavage/polyadenylation machinery thereby altering the cleavage/polyadenylation activity of the complex. This altering can be a reduction or increase in cleavage/polyadenylation activity.

The terms "polyadenylation complex", "cleavage/polyadenylation machinery" and "cleavage/polyadenylation complex" are used interchangeably herein.

4. BRIEF DESCRIPTION OF THE DRAWINGS

Fig. :1 shows elements of mammalian and yeast mRNA, respectively, which are involved in polyadenylation/cleavage of precursor mRNA

Fig. 2 shows a schematic representation of the gene targeting procedure. The TAP cassette is inserted at the C-terminus of a given yeast ORF by homologous recombination, generating the TAP-tagged fusion protein.

Fig. 3. shows the protein pattern obtained by separation of the members of the polyadenylation-complex of yeast using Pta1 as a bait using TAP. Protein bands for Cft1,

Cft2, Ysh1, Rna14, Pab1, Pcf11, Ref2, Pap1, Clp1, YKL059c, Pfs2, YGR156w, Fip1, Rna15, YKL018w, Glc7, Yth1, Ssu72, YOR179c and Pta1 (in bold) are labeled. (Further proteins identified as components of the yeast complex as described in the EXAMPLES-section (infra) are not stated in the figure)

Fig. 4 shows the protein pattern obtained by the separation of the members of the polyadenylation-complex in some of the reverse tagging-experiments and re-purification of a selection of the novel interactors. The baits using TAP used for the different experiments are given on top of each gel picture. The band constituting the protein used as the bait in the respective experiments is indicated by an arrow. Previously known members of the complex are listed in bold letters. (Note: only experiments using Cft1, Cft2, Pap1, Ref2, YKL059c, Pfs2, YOR179c and Pta1 as a bait are shown and only the proteins bands of Cft1, Cft2, Ysh1, Rna14, Pab1, Ref2, CLp1, Ygr156w, Fip1, Glc7, Yht1, Yor179c, Pta1, Pcf11, Pab1, Ykl059c, Pfs2, Rna15, Ykl018w and Ssu72 are labelled).

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to components of the cleavage/polyadenylation machinery of pre-cursor mRNA, the complete protein complex, uses of said components and complex as well as to methods of preparing same.

This is further described below. Also a description of the newly identified components of the cleavage/polyadenylation machinery is given below.

In more detail, the present invention relates to the following embodiments:

An isolated complex selected from complex (I) and comprising

(a) a first protein, or a functionally active fragment or functionally active derivative thereof, which first protein is selected from the group of proteins in Table 1, column A, or a mammalian homolog thereof, or a variant of said protein encoded by a nucleic acid that hybridizes to the nucleic acid of said protein or its complement under low stringency conditions; and

(b) a second protein, or a functionally active fragment or functionally active derivative thereof, which second protein is selected from the group of proteins in Table 1,

column B, or a mammalian homolog thereof, or a variant of said protein encoded by a nucleic acid that hybridizes to the nucleic acid of said protein or its complement under low stringency conditions, wherein said first protein and said second protein are members of a native cellular Polyadenylation-complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C, and a complex (II) comprising at least two second proteins.

The present invention further relates to a new protein complex which is useful for cleaving and/or polyadenylating a nucleic acid and which complex comprises at least one of the components according to the invention. Such a complex can be isolated from a natural source by applying the process according to the invention or can be reconstituted from the different components made available by the present invention.

Furthermore, the invention relates to an isolated complex comprising all proteins in column C of table 1, or the mammalian homologs of those proteins, or variants of said proteins encoded by nucleic acid that hybridises to the nucleic acid of any of said proteins or its complements under low stringency conditions, wherein proteins are members of a native cellular complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

Furthermore, the invention relates to an isolated complex that comprises all but 1,2,3,4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17,18,19,20,21,22,23,24,25,26,27 or 28 of all proteins in column C of table 1.

Furthermore, the invention relates to the complex as described above comprising a functionally active derivative of said first protein and/or a functionaliy active derivative of said second protein, wherein the functionally active derivative is a fusion protein

comprising said first protein or said second protein fused to an amino acid sequence different from the first protein or second protein, respectively.

The present invention further relates to a fusion protein comprising a component according to the invention. The fusion part, which can be added to the N-terminal, the C-terminal or into the amino acid sequence of the component according to the invention may comprise a few amino acids only e.g. at least five, which amino acids for example provide an epitope which is then be used as a target for affinity purification of the protein and the complex, respectively. Such a type of added amino acid is also termed "tag" throughout the present specification (optionally, the fusion protein may comprise even more than one such fusion partner).

In a preferred embodiment of the present invention, the protein components of the complex are vertebrate homologs of the yeast proteins, or a mixture of yeast and vertebrate homolog proteins. In a more preferred embodiment, the protein components of the complex are mammalian homologs of the yeast proteins, or a mixture of yeast and mammalian homolog proteins. In particular aspects, the native component proteins, or derivatives or fragments of the complex are obtained from a mammal such as mouse, rat, pig, cow, dog, monkey, human, sheep or horse. In another preferred embodiment, the protein components of the complex are human homologs of the yeast proteins, or a mixture of yeast and human homolog proteins. In yet another preferred embodiment, the protein components of the complex are a mixture of yeast, vertebrate, mammalian and/or human proteins.

The mammalian homologs or "orthologues" of the yeast proteins according to the invention can either be isolated based on the sequence homology of the yeast genes to the mammalian genes by cloning the respective gene applying conventional technology and expressing the protein from such gene or by isolating the mammalian proteins according to the process of the invention as explained in more detail below.

The derivatives of the proteins according to the invention can be produced e.g. by recombinant DNA technology applying the standard technology to modify the amino acid sequence of a given protein via the modification of the underlying gene using e.g. site directed mutagenesis, etc.

A protein that shows a certain degree of identity to the naturally occurring proteins from mammals and/or yeast, respectively, can also be prepared e.g. by applying recombinant DNA technology as described above for the derivatives according to the

invention. Alternatively, such protein can be isolated from natural sources by applying the process of the invention.

Furthermore, the invention relates to a complex as described above that is involved in the 3' end processing activity. Such a complex might also exist as a module or subcomplex of a larger physiological protein complex or assembly.

Furthermore, the invention relates to a complex as described above comprising a fragment of said first protein and/or a fragment of said second protein, which fragment binds to another protein component of said complex.

Furthermore, the invention relates to a complex as described above, wherein the functionally active derivative is a fusion protein comprising said first protein or said second protein preferentially fused to an affinity tag or label.

It is further directed to complexes comprising a fusion protein which comprises a component of the complex or a fragment thereof linked via a covalent bond to an amino acid sequence different from said component protein, as well as nucleic acids encoding the protein, fusions and fragments thereof. For example, the non-component protein portion of the fusion protein, which can be added to the N-terminal, the C-terminal or inserted into the amino acid sequence of the complex component can comprise a few amino acids, which provide an epitope that is used as a target for affinity purification of the fusion protein and/or complex.

Furthermore the invention relates to a process for processing RNA comprising the step of bringing into contact any of the complexes described above with RNA, such that RNA is processed.

Furthermore, the invention relates to an antibody or a fragment of said antibody containing the binding domain thereof, which binds the complex as described above of claim and which does not bind the first protein when uncomplexed or the second protein when uncomplexed.

Furthermore, the invention relates to a pharmaceutical composition comprising the protein complex as described above and a pharmaceutically acceptable carrier.

Moreover, the present invention provides a process for the identification and/or preparation of an effector of a composition according to the invention which process comprises the steps of bringing into contact the composition of the invention or of a component thereof with a compound, a mixture of compounds or a library of compounds and determining whether the compounds or certain compounds of the mixture or library bind to the composition of the invention and/or a component thereof and/or affects the

biological activity of such a composition or component and then optionally further purifying the compound positively tested as effector by such a process.

A major application of the composition according to the invention results in the identification of an active agent capable of binding thereto. Hence, the compositions of the invention are useful tools in screening for new pharmaceutical drugs.

Furthermore, the invention relates to a method for screening for a molecule that modulates directly or indirectly the function, activity, composition or formation of the complex as described above comprising the steps of :

- (a) exposing said complex, or a cell or organism containing said complex to one or more candidate molecules; and
- (b) determining the amount of, the 3' end processing activity for mRNA of, or protein components of, said complex, wherein a change in said amount, activity, or protein components relative to said amount, activity or protein components in the absence of said candidate molecules indicates that the molecules modulate function, activity or composition of said complex.

Furthermore, the invention relates to a method as described above, wherein the amount of said complex is determined.

Furthermore, the invention relates to a method as described above, wherein the activity of said complex is determined.

Furthermore, the invention relates to a method as described above, wherein said determining step comprises isolating from the cell or organism said complex to produce said isolated complex and contacting said isolated complex with the substrate under conditions conducive to binding to the complex.

Furthermore, the invention relates to a method as described above, wherein the protein components of said complex are determined.

Furthermore, the invention relates to a method as described above, wherein said determining step comprises determining whether any of the proteins listed in column B of table 1 of said complex or the mammalian homologs thereof, or variant of said proteins encoded by a nucleic acid that hybridises to the nucleic acids of any of said proteins or its complements under low stringency conditions, is present in the complex, wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM

EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

Furthermore, the invention relates to a method as described above, wherein said method is a method of screening for a drug for treatment or prevention of diseases and disorders, preferably diseases or disorders such as infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis and cancer.

Furthermore, the invention relates to a method for screening for a molecule that binds the complex as described above comprising the following steps:

- (a) exposing said complex, or a cell or organism containing said complex, to one or more candidate molecules; and
- (b) determining whether said complex is bound by any of said candidate molecules.

Furthermore, the invention relates to a method for diagnosing or screening for the presence of a disease or disorder or a predisposition for developing a disease or disorder in a subject, which disease or disorder is characterized by an aberrant amount of, the 3' end processing activity for mRNA biochemical activity of, or component composition or formation of, the complex as described above, comprising determining the amount of, the 3' end processing activity for mRNA of, or protein components of, said complex in a sample derived from a subject, wherein a difference in said amount, activity, or protein components of, said complex in an analogous sample from a subject not having the disease or disorder or predisposition indicates the presence in the subject of the disease or disorder or predisposition.

Furthermore, the invention relates to a method as described above, wherein the amount of said complex is determined.

Furthermore, the invention relates to a method as described above, wherein the activity of said complex is determined.

Furthermore, the invention relates to a method as described above, wherein said determining step comprises isolating from the cell or organism said complex to produce said isolated complex and contacting said isolating complex with the substrate under conditions conducive to binding to the complex.

Furthermore, the invention relates to a method as described above, wherein the protein components of said complex are determined.

Furthermore, the invention relates to a method as described above, wherein said determining step comprises determining whether any of the proteins listed in column B of table 1 of said complex or the mammalian homologs thereof, or variant of said proteins encoded by a nucleic acid that hybridises to the nucleic acids of any of said proteins or its complements under low stringency conditions, is present in the complex, wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

Furthermore, the invention relates to a method for treating or preventing a disease or disorder characterized by an aberrant amount of, the 3' end processing activity for mRNA of, or component composition or formation of, the complex as described above, comprising administering to a subject in need of such treatment or prevention a therapeutically effective amount of one or more molecules that modulate the amount of, the 3' end processing activity for mRNA of, or protein components or formation of, said complex.

Furthermore, the invention relates to a method as described above, wherein said disease or disorder involves decreased levels of the amount or activity of said complex. Furthermore, the invention relates to a method as described above, wherein said disease or disorder involves increased levels of the amount or activity of said complex.

Furthermore, the invention relates to the use of a molecule that modulates the amount of, the 3' end processing activity for mRNA of, or protein components or formation of the complex as described above for the manufacture of a medicament for the treatment or prevention of a disease or disorder, preferably diseases or disorders such as infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis; cancer

The present invention further relates to the use of the products according to the invention in therapy wherein the products according to the invention are useful as a target for a therapeutic drug. It is known from the prior art that mRNA 3'-end processing is involved in viral growth, in the development of cancer and in certain neurodegenerative diseases. By having identified new components of the cleavage/polyadenylation

machinery the present invention, hence, offers new targets for treating viral diseases, cancer and neurodegenerative diseases. By affecting the biological activity of the components of the invention and/or by affecting the complex as a whole the cleavage/polyadenylation activity thereof can be influenced depending by the needs of the patient to be treated.

The present invention further relates to a pharmaceutical composition comprising a product according to the invention. Such pharmaceutical composition contains beside the product according to the invention as active ingredient further excipients and additives as known by a skilled person. The present invention, hence, allows the identification of new effectors which affect the biological activity of the cleavage/polyadenylation machinery of precursor RNA. Said effectors than can be used to modify the cleavage/polyadenylation machinery in a cell by introducing an effector into a cell. Moreover, the mRNA processing activity of a given cell can also be affected by introducing a product according to the invention into such cell.

Furthermore, the invention relates to a kit comprising in one or more containers

- (a) an isolated first protein, or a functionally active fragment or functionally active derivative thereof selected from the proteins in column A of table 1 of a given complex or a mammalian homolog thereof, or a variant of said protein encoded by a nucleic acid that hybridises to the nucleic acid of said protein or its complement under low stringency conditions; and
- (b) an isolated second protein, or a functionally active fragment or functionally active derivative thereof selected from the proteins in column B of table 1 of a given complex or a mammalian homolog thereof, or a variant of said protein encoded by a nucleic acid that hybridises to the nucleic acid of said protein or its complement under low stringency conditions, wherein said first and said second protein are members of a native cellular complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

Furthermore, the invention relates to a kit comprising in a container the isolated complex as described above or the antibody as described above, optionally together with

further reagents and working instructions. The further reagents may be, for example, buffers, substrates for enzymes but also carrier material such as beads, filters, microarrays and other solid carries. The working instructions may indicate how to use the ingredients of the kit in order to perform a desired assay..

Furthermore, the invention relates to such kits for use in processing of RNA and for use in the diagnosis, prognosis and screening in or for the diseases mentioned above.

The present invention further relates to a kit for processing RNA which kit comprises a product according to the invention. Such a kit may contain e.g. expression vectors encoding the essential components of the cleavage/polyadenylation machinery which components after being expressed can be reconstituted in order to form a biologically active cleavage/polyadenylation complex. Such a kit preferably also contains the required buffers and reagents together with the working instructions.

The present invention further relates to a kit for the diagnosis of diseases of mammals which kit comprises a product according to the invention. As stated above the polyadenylation/cleavage machinery is involved in a large number of diseases. If said machinery activity is changed due to e.g. mutations of some components thereof and/or effectors, this may have severe implications on the affected organism. The kit according to the present invention containing the products according to the invention allows to examine as to whether the cleavage/polyadenylation machinery in a given sample might show some defects. Such a kit may be used to determine genetic defects of the genes encoding the components of the cleavage/polyadenylation machinery.

Furthermore, the invention relates to a complex as described above, or the antibody or fragment as described above, for use in a method of diagnosing a disease or disorder, preferably the diseases or disorders as mentioned above.

Furthermore, the invention relates to a method for the production of a pharmaceutical composition comprising carrying out the method as described above to identify a molecule that modulates the function, activity or formation of said complex, and further comprising mixing the identified molecule with a pharmaceutically acceptable carrier.

Furthermore, the invention relates to a process for preparing complex as described above and optionally the components thereof comprising the following steps: expressing such a protein in a target cell, isolating the protein complex which is attached

to the tagged protein, and optionally disassociating the protein complex and isolating the individual complex members.

Furthermore, the invention relates to the process as described above characterized in that the tagged protein comprises two different tags which allow two separate affinity purification steps.

Furthermore, the invention relates to the process as described above, characterized in that two tags are separated by a cleavage site for a protease.

Furthermore, the invention relates to a component of the said complex obtainable by a process as described above.

The present invention further relates to a composition, preferably a protein complex, which is obtainable by the method comprising the following steps: tagging a protein as defined above, i.e. a protein which forms part of a protein complex, with a moiety, preferably an amino acid sequence, that allows affinity purification of the tagged protein and expressing such protein in a target cell and isolating the protein complex which is attached to the tagged protein. The details of such purification are described in WO 00/09716 and in Rigaut, G. et al. (1999), Nature Biotechnology, Vol. 17 (10): 1030-1032 and further herein below. The tagging can essentially be performed with any moiety which is capable of providing a specific interaction with a further moiety, e.g. in the sense of a ligand receptor interaction, antigen antibody interaction or the like. The tagged protein can also be expressed in an amount in the target cell which comes close to the physiological concentration in order to avoid a complex formation merely due to high concentration of the expressed protein but not reflecting the natural occurring complex.

In a further preferred embodiment, the composition is obtained by using a tagged protein which comprises two different tags which allow two different affinity purification steps. This measure allows a higher degree of purification of the composition in question. In a further preferred embodiment the tagged protein comprises two tags that are separated by a cleavage site for a protease. This allows a step-by-step purification on affinity columns.

Furthermore, the invention relates to a complex as described above and/or protein thereof as a target for an active agent of a pharmaceutical, preferably a drug target in the treatment or prevention of disease or disorder, preferably diseases or disorders as mentioned above..

Furthermore, the invention relates to the Ycl046w (SEQ ID: 59), Ygr156w (SEQ ID: 61), Yhl035c (SEQ ID:63), Ykl018w (SEQ ID:179), Ylr221c (SEQ ID: 67), Yml030w (SEQ ID:69), and Yor17c (SEQ ID:71), the mammalian homologs/orthologs of said proteins and functionally active fragments and derivatives of said proteins and the mammalian homologs thereof carrying one or more amino acid substitutions, deletions and/or additions and the nucleic acid encoding said proteins or said homologs, orthologs and functionally active fragments and derivatives thereof.

Such a nucleic acid may be used for example to express a desired tagged protein in a given cell for the isolation of a complex or component according to the invention. Such a nucleic acid may also be used for the identification and isolation of genes from other organisms by cross species hybridization.

The component according to the invention is preferably a protein component which can be further modified e.g. by carbohydrate residues. The components can either be prepared by recombinant DNA technology based on the sequences provided by the present invention or can be isolated from a biological source by using the process according to the invention.

The present invention further relates to a fusion protein comprising a component according to the invention. The fusion part, which can be added to the N-terminal, the C-terminal or into the amino acid sequence of the component according to the invention may comprise a few amino acids only e.g. at least five, which amino acids for example provide an epitope which is then be used as a target for affinity purification of the protein and the complex, respectively. Such a type of added amino acid is also termed "tag" throughout the present specification (optionally, the fusion protein may comprise even more than one such fusion partner).

The present invention further relates to a construct, preferably a vector construct, which comprises a nucleic acid as described above. Such constructs may comprise expression controlling elements such as promoters, enhancers and terminators in order to express the nucleic acids in a given host cell, preferably under conditions which resemble the physiological concentrations.

The present invention further relates to a construct which comprises the nucleic acid according to the invention and at least one further nucleic acid which is normally not associated with the nucleic acid according to the invention. Such a construct is preferably a vector which preferably is capable of replicating in a given cell and contains the necessary transcription control elements for expressing the nucleic acid according to the

invention in a given expression system. Moreover, such vector construct may contain selection markers.

The present invention further relates to a host cell containing a construct as defined above.

Such a host cell can be, e.g., any eukaryotic cell such as yeast, plant or mammalian, whereas human cells are preferred. Such host cells may form the starting material for isolation of a complex according to the present invention.

The present invention also relates to a host cell containing a nucleic acid according to the invention or a construct according to the invention. Such a host cell may contain an expression vector which encodes a component according to the invention which component may serve as a bait in order to isolate the further proteins of the complex and which at least partly interact with the bait. Host cells can be prokaryotic and eukaryotic cells, whereas mammalian host cells are preferred.

Animal models and methods of screening for modulators (i.e., agonists, and antagonists) of the amount of, activity of, or protein component composition of, a complex of the present invention are also provided.

Below is a more detailed list of the newly identified components of the polyadenylation complex (see also Tab. 1). The Accession-Number stated is the GenBank-Accession number for the protein.

Act1: Is a known and essential protein (GenBank Acc. No. BAA21512.1), which has been shown to be involved in Pol II transcription and has been found to be associated with histone acetylation. It serves as a structural protein.

Cka1: Is a known and non-essential protein (GenBank Acc. No. CAA86916.1), which has been found to be involved in Polymerase III transcription and has been found to be associated with the Casein kinase II complex.

Eft2: The translation elongation factor EF-2 is a known protein involved in protein synthesis (GenBank AAB64827.1)

Eno2: Is a known and essential protein (GenBank Acc. No. AAB68019.1). It has been shown to have lyase activity and is known to be involved in carbohydrate metabolism.

Glc7 (YER133w) is also a known protein (GenBank Acc. No. AAC03231.1). It is also an essential protein and is a Type I protein serine threonine phosphatase which has been implicated in distinct cellular roles, such as carbohydrate metabolism, meiosis, mitosis and cell polarity. Its occurrence in the cleavage/polyadenylation machinery has not been known before.

Gpm1: This protein is a phosphoglycerate mutase that converts 2-phosphoglycerate to 3-phosphoglycerate in glycolysis. It is an essential protein (GenBank: CAA81994.1)

Hhf2: Is a known and non-essential protein (GenBank Acc. No. CAA95892.1) which has been shown to be involved in DNA-binding. It has previously been linked to Histone octamer and the RNA polymerase I upstream activation factor.

Hta1: Is a known and non-essential protein (GenBank Acc. No. CAA88505.1) which has DNA-binding capability and has been shown to be involved in polymerase II transcription.

Hsc82: Is a non-essential protein so far being associated with protein folding. (GenBank Acc. No: CAA89919.1)

Imd2: Is an Inosine-5'-monophosphate dehydrogenase so far being associated with nucleotide metabolism. It is non-essential. (GenBank Acc.-No.: AAB69728.1)

Imd4: Is a non-essential protein with similarity to Imd2 so far being associated with nucleotide metabolism (GenBank Acc-No.: CAA86719.1)

Met6: Is a homocysteine methyltransferase so far being associated with amino-acid metabolism (GenBank Acc.-No.: AAB64646.1)

Pdc1: Is a pyruvate decarboxylase isozyme1 so far being associated with carbohydrate metabolism (GenBank Acc.-No.: CAA97573.1)

Pfk1: Is a known protein (GenBank Acc. No. CAA97268.1) which has previously been described as part of the phosphofructokinase complex.

Ref2 (YDR195w) is a known protein (GenBank Acc. No. CAA88708.1). It is a non-essential gene product. It has been shown to be involved in 3'-end formation prior to the final polyadenylation step. However, Ref2 has never been identified before as a component of the 3'-end processing machinery. Ref2 has been shown to interact with Glc7, another new component of the cleavage/polyadenylation machinery.

Sec13: Is a known and essential protein (GenBank Acc. No. AAB67426.1).

Sec31: Is a known and essential protein (GenBank Acc. No. CAA98772.1)

Ssa3: Is a known and non-essential protein (GenBank Acc. No. CAA84896.1) which so far has been implicated with protein folding/protein transport.

Ssu72 (YNL222w) is also a known protein (GenBank Acc. No. CAA96125.1) and is an essential phylogenetically conserved protein which has been shown to interact with the general transcription factor TFIIB (Sua7). TFIIB is an essential component of the RNA polymerase II (RNAP II) core transcriptional machinery. It is thought that this interaction plays a role in the mechanism of start site selection by RNAP II. The finding according to the present invention that Ssu72 is associated with Pta1 is likely to be relevant since it is believed that mRNA 3'-end formation is linked with other nuclear processes like transcription, capping and splicing. Furthermore, Ssu 72 has also been clearly identified in a "reverse tagging experiment" as explained herein below by using some of the Pta1 associated proteins as bait. However, when Ssu72 itself was used as a bait associated proteins were not found most likely due to the fact that the addition of a C-terminal tag renders Ssu72 non-functional.

Taf60: Is a known and essential protein (GenBank Acc. No. CAA96819.1) which has been shown to be involved in Polymerase II transcription.

Tkl1: Is a non-essential transketolase so far being associated with amino-acid metabolism and carbohydrate metabolism (GenBank Acc-No.: CAA89191.1)

Tsa1: Translation initiation factor eIF5 which so far has been shown to catalyze hydrolysis of GTP on the 40S ribosomal subunit-initiation complex followed by joining to 60S ribosomal subunit. (GenBank Acc.-No.: CAA92145.1)

Tye7: Is a known protein (GenBank Acc. No. CAA99671.1). It has been shown to be a basic helix-loop-helix transcription factor.

Vid24: Is a known and non-essential protein (GenBank Acc. No. CAA89320.1) which has previously been associated with protein degradation and vesicular transport.

Vps53: Is a known protein (GenBank Acc. No. CAA89320.1) which has been found to play a role in protein sorting.

YCL046w: Is a non-essential protein (GenBank Acc. No. CAA42371.1).

YGR156w is the protein product of an essential gene. This protein also contains a RNA binding motif. (GenBank Acc. No. CAA97170.1).

YHL035c: Is a known and non-essential protein (GenBank Acc. No. AAB65047.1). It is a member of the ATP-binding cassette superfamily.

YKL018w is also an essential protein containing a WD40 domain which is a typical protein binding domain. (GenBank Acc. No. CAA81853.1)

YLR221c: Is a protein of unknown function (GenBank Acc. No. AAB67410.1)

YML030w: Is a protein of unknown function (GenBank Acc. No. CAA86625.1)

YOR179c shows significant sequence similarity to Ysh1 (GenBank Acc. No. CAA99388.1)

Two further proteins for which binary interactions with members of the polyadenylation complex as known so far have been shown before have also been purified with the complex:

YKL059c: is the product of an essential gene and is a zinc binding protein containing a C2HC Zinc finger. The presence of this domain predicts a RNA binding function of YKL059c. We believe the corresponding gene product is identical to Pfs1, a protein which has been mentioned in several publications, but which has never been annotated in the databases (for review see Keller, W. and Minvielle-Sebastia (1997). *Curr Opin Cell Biol* 11: 352-357). (GenBank Acc. No. CAA81896.1)

Tif4632: Is a known and non-essential protein (GenBank Acc. No. CAA96751.1) which has been shown to have an RNA-binding/translation factor activity and is involved in protein synthesis.

TABLES:

Table 1: Composition of the Complex (Cleavage/polyadenylation machinery):

First column ('Entry point') lists the bait proteins (TAP-tag fusion proteins) that have been chosen for the isolation of the given complex. Note: in several cases, different baits have been used for validation in reverse tagging experiments.

Second column ('Interactions') briefly lists any known interactions between different members of the complex (Abbreviations: '2-hybrid': interaction as identified in yeast-2-hybrid screens; 'far-western': interaction as identified in far-western experiments; 'coipp': interaction as identified by co-immunoprecipitation experiments; 'high-throughput 2 hybrid': interaction as identified by high-throughput yeast-2-hybrid screens; 'copurification': interaction as identified by copurification experiments; 'immuno-affinity-columns': interaction as identified in experiments using immuno-affinity columns; 'in vitro binding': interaction as identified in in-vitro-binding experiments. If a core complex has been known previously containing several of the identified proteins, the name of the complex is stated.

Third column ('Proteins found') lists all proteins which have been identified in the particular complex.

Fourth column ('COLUMN A, 'Known components of the complex') lists the components of the complex as found by Cellzome, which have been known to interact with other members of the complex as identified herein. (see also third column).

Fifth column ('COLUMN B, 'Novel proteins') lists the novel members of the complex as provided in the invention.

Sixth column ('Column C, cleavage/polyadenylation machinery'): lists again all components of the cleavage/polyadenylation machinery as identified herein

Seventh column (COLUMN C, 'Activity of the complex'): List the biochemical activities of the newly identified complex.

Eighth column (COLUMN D, 'Proteins of unknown function'): Separately lists again the members of the newly identified complex which previously have not been annotated.

Ninth column ('localization') indicates the localization of the identified complex (Abbreviations: c: cytoplasm; b: membrane; e: ER/Golgi/vesicles; m: mitochondria; n: nucleus; u: unknown)

Table 2: Individual Yeast Proteins of the Complexes

A) Table lists in alphabetical order all yeast proteins which have been identified as members of the complex presented herein. Furthermore, the SEQ ID of the proteins are listed as used herein. Further columns lists the Accession-Number of the respective sequences in MIPS, SWISS-PROT, SGD and Genbank. In addition, where applicable, the GenBank accession numbers of the respective orthologues in humans, *C.elegans* and *Drosophila* are listed.

B) Table lists again the proteins and SEQ ID as in part A. In addition, the table contains an overview about what has been previously reported on the protein, the biochemical function thereof and the cellular function thereof as stated in YPD (Constanzo, M.C. et al., 2001, Nucl. Acid Res, 29: 75-9; Hodges, P.E. et al., 1999, Nucl. Acids Res 27: 69-73).

Table 3: Medical Application of the Complex:

First column ('Name of complex') lists again the name of the complex as used herein.

Second column ('Cellular role') lists keyword on the cellular role of the complex

Third column ('Medical applications') lists disorder, diseases, disease areas etc. which are treatable and/or preventable and/or diagnosable etc. by therapeutics and methods interacting with/acting via the complex.

Table 4: Characterization of previously undescribed individual proteins of the complexes: The table provides data on proteins which have not been annotated previously but which have now been linked to a functional complex as described in table 2. Names are listed on the left. In addition the table contains a list of motifs found by sequence analysis which has been part of the invention provided herein. Furthermore, the predicted known human orthologues are listed on the right (By SWISS-PROT Accession numbers). Used Abbreviations are listed at the end of the table. The function of the individual proteins as deduced from the association with the complex, the sequence analysis and the analysis of the predicted orthologues is listed in the second column ('Putative function').

Table5: Overview on Experimental Steps: The tables illustrates the construction of a yeast strain expressing a TAP-tagged bait in a high-throughput fashion.

Table 6: Known and Novel Components of the yeast mRNA 3'-end processing machinery (the cleavage/polyadenylation complex): Top part of the table states the different known subcomponents of the polyadenylation complex, the function thereof, the proteins constituting the different subcomplexes as known so far (including their molecular weight and sequence motifs contained in the protein). Bottom part lists the novel components of the complex as provided herein

5.1. PROTEIN COMPLEXES

The protein complexes of the present invention and their component proteins are described in the Tables 1,2,3,4,6 (whereas Table 6 gives an overview on the construction of the yeast strains). The protein complexes and component proteins can be obtained by methods well known in the art for protein purification and recombinant protein expression. For example, the protein complexes of the present invention can be isolated using the TAP method described in Section 6, *infra*, and in WO 00/09716 and Rigaut et al., 1999, Nature Biotechnology 17:1030-1032, which are each incorporated by reference in their entirety. Additionally, the protein complexes can be isolated by immunoprecipitation of the component proteins and combining the immunoprecipitated proteins. The protein complexes can also be produced by recombinantly expressing the component proteins and combining the expressed proteins.

The nucleic and amino acid sequences of the component proteins of the protein complexes of the present invention are provided herein (SEQ ID NOS:1-2670), and can be obtained by any method known in the art, *e.g.*, by PCR amplification using synthetic primers hybridizable to the 3' and 5' ends of each sequence, and/or by cloning from a cDNA or genomic library using an oligonucleotide specific for each nucleotide sequence.

Homologs (*e.g.*, nucleic acids encoding component proteins from other species) or other related sequences (*e.g.*, variants, paralogs) which are members of a native cellular protein complex can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular nucleic acid sequence as a probe, using methods well known in the art for nucleic acid hybridization and cloning.

Exemplary moderately stringent hybridization conditions are as follows: prehybridization of filters containing DNA is carried out for 8 hours to overnight at 65°C in buffer composed of 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 µg/ml denatured salmon sperm DNA. Filters are hybridized for 48 hours at 65 °C in prehybridization mixture containing 100 µg/ml denatured salmon sperm DNA and 5-20 X 10⁶ cpm of ³²P-labeled probe. Washing of filters is done at 37 °C for 1 hour in a solution containing 2X SSC, 0.01% PVP, 0.01% Ficoll, and 0.01% BSA. This is followed by a wash in 0.1X SSC at 50 °C for 45 min before autoradiography. Alternatively, exemplary conditions of high stringency are as follows: *e.g.*, hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65 °C, and washing in 0.1xSSC/0.1% SDS at 68 °C (Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3). Other conditions of high stringency which may be used are well known in the art. Exemplary low stringency hybridization conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 µg/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

For recombinant expression of one or more of the proteins, the nucleic acid containing all or a portion of the nucleotide sequence encoding the protein can be inserted into an appropriate expression vector, *i.e.*, a vector that contains the necessary

elements for the transcription and translation of the inserted protein coding sequence. The necessary transcriptional and translational signals can also be supplied by the native promoter of the component protein gene, and/or flanking regions.

A variety of host-vector systems may be utilized to express the protein coding sequence. These include but are not limited to mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); microorganisms such as yeast containing yeast vectors; or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA. The expression elements of vectors vary in their strengths and specificities. Depending on the host-vector system utilized, any one of a number of suitable transcription and translation elements may be used.

In a preferred embodiment, a complex of the present invention is obtained by expressing the entire coding sequences of the component proteins in the same cell, either under the control of the same promoter or separate promoters. In yet another embodiment, a derivative, fragment or homolog of a component protein is recombinantly expressed. Preferably the derivative, fragment or homolog of the protein forms a complex with the other components of the complex, and more preferably forms a complex that binds to an anti-complex antibody.

The present invention further relates to an antibody which reacts with a product according to the invention. Such an antibody might be used e.g. during purification of the machinery from a given source by affinity purification methods. Moreover, the antibody might be used in diagnosis in order to detect changes and/or modifications of a product according to the invention in a given sample.

Any method available in the art can be used for the insertion of DNA fragments into a vector to construct expression vectors containing a chimeric gene consisting of appropriate transcriptional/translational control signals and protein coding sequences. These methods may include *in vitro* recombinant DNA and synthetic techniques and *in vivo* recombinant techniques (genetic recombination). Expression of nucleic acid sequences encoding a component protein, or a derivative, fragment or homolog thereof, may be regulated by a second nucleic acid sequence so that the gene or fragment thereof is expressed in a host transformed with the recombinant DNA molecule(s). For example, expression of the proteins may be controlled by any promoter/enhancer known in the art. In a specific embodiment, the promoter is not native to the gene for the

component protein. Promoters that may be used can be selected from among the many known in the art, and are chosen so as to be operative in the selected host cell.

In a specific embodiment, a vector is used that comprises a promoter operably linked to nucleic acid sequences encoding a component protein, or a fragment, derivative or homolog thereof, one or more origins of replication, and optionally, one or more selectable markers (e.g., an antibiotic resistance gene).

In another specific embodiment, an expression vector containing the coding sequence, or a portion thereof, of a component protein, either together or separately, is made by subcloning the gene sequences into the EcoRI restriction site of each of the three pGEX vectors (glutathione S-transferase expression vectors; Smith and Johnson, 1988, Gene 7:31-40). This allows for the expression of products in the correct reading frame.

Expression vectors containing the sequences of interest can be identified by three general approaches: (a) nucleic acid hybridization, (b) presence or absence of "marker" gene function, and (c) expression of the inserted sequences. In the first approach, coding sequences can be detected by nucleic acid hybridization to probes comprising sequences homologous and complementary to the inserted sequences. In the second approach, the recombinant vector/host system can be identified and selected based upon the presence or absence of certain "marker" functions (e.g., resistance to antibiotics, occlusion body formation in baculovirus, etc.) caused by insertion of the sequences of interest in the vector. For example, if a component protein gene, or portion thereof, is inserted within the marker gene sequence of the vector, recombinants containing the encoded protein or portion will be identified by the absence of the marker gene function (e.g., loss of beta-galactosidase activity). In the third approach, recombinant expression vectors can be identified by assaying for the component protein expressed by the recombinant vector. Such assays can be based, for example, on the physical or functional properties of the interacting species in *in vitro* assay systems, e.g., formation of a complex comprising the protein or binding to an anti-complex antibody.

Once recombinant component protein molecules are identified and the complexes or individual proteins isolated, several methods known in the art can be used to propagate them. Using a suitable host system and growth conditions, recombinant expression vectors can be propagated and amplified in quantity. As previously described, the expression vectors or derivatives which can be used include, but are not limited to, human or animal viruses such as vaccinia virus or adenovirus; insect viruses

such as baculovirus, yeast vectors; bacteriophage vectors such as lambda phage; and plasmid and cosmid vectors.

In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies or processes the expressed proteins in the specific fashion desired. Expression from certain promoters can be elevated in the presence of certain inducers; thus expression of the genetically-engineered component proteins may be controlled. Furthermore, different host cells have characteristic and specific mechanisms for the translational and post-translational processing and modification (e.g., glycosylation, phosphorylation, etc.) of proteins. Appropriate cell lines or host systems can be chosen to ensure that the desired modification and processing of the foreign protein is achieved. For example, expression in a bacterial system can be used to produce an unglycosylated core protein, while expression in mammalian cells ensures "native" glycosylation of a heterologous protein. Furthermore, different vector/host expression systems may effect processing reactions to different extents.

In other specific embodiments, a component protein or a fragment, homolog or derivative thereof, may be expressed as fusion or chimeric protein product comprising the protein, fragment, homolog, or derivative joined via a peptide bond to a heterologous protein sequence of a different protein. Such chimeric products can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acids to each other by methods known in the art, in the proper coding frame, and expressing the chimeric products in a suitable host by methods commonly known in the art. Alternatively, such a chimeric product can be made by protein synthetic techniques, e.g., by use of a peptide synthesizer. Chimeric genes comprising a portion of a component protein fused to any heterologous protein-encoding sequences may be constructed.

In particular, protein component derivatives can be made by altering their sequences by substitutions, additions or deletions that provide for functionally equivalent molecules. Due to the degeneracy of nucleotide coding sequences, other DNA sequences that encode substantially the same amino acid sequence as a component gene or cDNA can be used in the practice of the present invention. These include but are not limited to nucleotide sequences comprising all or portions of the component protein gene that are altered by the substitution of different codons that encode a functionally equivalent amino acid residue within the sequence, thus producing a silent change. Likewise, the derivatives of the invention include, but are not limited to, those containing, as a primary amino acid sequence, all or part of the amino acid sequence of

a component protein, including altered sequences in which functionally equivalent amino acid residues are substituted for residues within the sequence resulting in a silent change. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity that acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

In a specific embodiment, up to 1%, 2%, 5%, 10%, 15% or 20% of the total number of amino acids in the wild type protein are substituted or deleted; or 1, 2, 3, 4, 5, or 6 amino acids are inserted, substituted or deleted relative to the wild type protein.

In a specific embodiment of the invention, the nucleic acids encoding a protein component and protein components consisting of or comprising a fragment of or consisting of at least 6 (continuous) amino acids of the protein are provided. In other embodiments, the fragment consists of at least 10, 20, 30, 40, or 50 amino acids of the component protein. In specific embodiments, such fragments are not larger than 35, 100 or 200 amino acids. Derivatives or analogs of component proteins include, but are not limited, to molecules comprising regions that are substantially homologous to the component proteins, in various embodiments, by at least 30%, 40%, 50%, 60%, 70%, 80%, 90% or 95% identity over an amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to a sequence encoding the component protein under stringent, moderately stringent, or nonstringent conditions.

The protein component derivatives and analogs of the invention can be produced by various methods known in the art. The manipulations which result in their production can occur at the gene or protein level. For example, the cloned gene sequences can be modified by any of numerous strategies known in the art (Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). The sequences can be cleaved at appropriate sites with

restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated *in vitro*. In the production of the gene encoding a derivative, homolog or analog of a component protein, care should be taken to ensure that the modified gene retains the original translational reading frame, uninterrupted by translational stop signals, in the gene region where the desired activity is encoded.

Additionally, the encoding nucleic acid sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy pre-existing ones, to facilitate further *in vitro* modification. Any technique for mutagenesis known in the art can be used, including but not limited to, chemical mutagenesis and *in vitro* site-directed mutagenesis (Hutchinson et al., 1978, J. Biol. Chem 253:6551-6558), amplification with PCR primers containing a mutation, etc.

Once a recombinant cell expressing a component protein, or fragment or derivative thereof, is identified, the individual gene product or complex can be isolated and analyzed. This is achieved by assays based on the physical and/or functional properties of the protein or complex, including, but not limited to, radioactive labeling of the product followed by analysis by gel electrophoresis, immunoassay, cross-linking to marker-labeled product, etc.

The component proteins and complexes may be isolated and purified by standard methods known in the art (either from natural sources or recombinant host cells expressing the complexes or proteins), including but not restricted to column chromatography (*e.g.*, ion exchange, affinity, gel exclusion, reversed-phase high pressure, fast protein liquid, etc.), differential centrifugation, differential solubility, or by any other standard technique used for the purification of proteins. Functional properties may be evaluated using any suitable assay known in the art.

Alternatively, once a component protein or its derivative, is identified, the amino acid sequence of the protein can be deduced from the nucleic acid sequence of the chimeric gene from which it was encoded. As a result, the protein or its derivative can be synthesized by standard chemical methods known in the art (*e.g.*, Hunkapiller et al., 1984, Nature 310: 105-111).

Manipulations of component protein sequences may be made at the protein level. Included within the scope of the invention is a complex in which the component proteins or derivatives and analogs that are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known

protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH₄, acetylation, formylation, oxidation, reduction, metabolic synthesis in the presence of tunicamycin, etc.

In specific embodiments, the amino acid sequences are modified to include a fluorescent label. In another specific embodiment, the protein sequences are modified to have a heterofunctional reagent; such heterofunctional reagents can be used to crosslink the members of the complex.

In addition, complexes of analogs and derivatives of component proteins can be chemically synthesized. For example, a peptide corresponding to a portion of a component protein, which comprises the desired domain or mediates the desired activity *in vitro* (e.g., complex formation) can be synthesized by use of a peptide synthesizer. Furthermore, if desired, non-classical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the protein sequence.

In cases where natural products are suspected of being mutant or are isolated from new species, the amino acid sequence of a component protein isolated from the natural source, as well as those expressed *in vitro*, or from synthesized expression vectors *in vivo* or *in vitro*, can be determined from analysis of the DNA sequence, or alternatively, by direct sequencing of the isolated protein. Such analysis can be performed by manual sequencing or through use of an automated amino acid sequenator.

The complexes can also be analyzed by hydrophilicity analysis (Hopp and Woods, 1981, Proc. Natl. Acad. Sci. USA 78:3824-3828). A hydrophilicity profile can be used to identify the hydrophobic and hydrophilic regions of the proteins, and help predict their orientation in designing substrates for experimental manipulation, such as in binding experiments, antibody synthesis, etc. Secondary structural analysis can also be done to identify regions of the component proteins, or their derivatives, that assume specific structures (Chou and Fasman, 1974, Biochemistry 13:222-23). Manipulation, translation, secondary structure prediction, hydrophilicity and hydrophobicity profile predictions, open reading frame prediction and plotting, and determination of sequence homologies, etc., can be accomplished using computer software programs available in the art.

Other methods of structural analysis including but not limited to X-ray crystallography (Engstrom, 1974 Biochem. Exp. Biol. 11:7-13), mass spectroscopy and

gas chromatography (Methods in Protein Science, J. Wiley and Sons, New York, 1997), and computer modeling (Fletterick and Zoller, eds., 1986, Computer Graphics and Molecular Modeling, In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York) can also be employed.

5.2. ANTIBODIES TO PROTEIN COMPLEXES

According to the present invention, a protein complex of the present invention comprising a first protein, or a functionally active fragment or functionally active derivative thereof, selected from the group consisting of proteins listed in column A of table 1; and a second protein, or a functionally active fragment or functionally active derivative thereof, selected from the group consisting of proteins listed in column B of table 1, or a functionally active fragment or functionally active derivative thereof, can be used as an immunogen to generate antibodies which immunospecifically bind such immunogen. According to the present invention, also a protein complex of the present invention can be used as an immunogen to generate antibodies which immunospecifically bind to such immunogen comprising all proteins listed in column C of table 1

Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library. In a specific embodiment, antibodies to a complex comprising human protein components are produced. In another embodiment, a complex formed from a fragment of said first protein and a fragment of said second protein, which fragments contain the protein domain that interacts with the other member of the complex, are used as an immunogen for antibody production. In a preferred embodiment, the antibody specific for the complex in that the antibody does not bind the individual protein components of the complex.

Polyclonal antibodies can be prepared as described above by immunizing a suitable subject with a polypeptide of the invention as an immunogen. Preferred polyclonal antibody compositions are ones that have been selected for antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred polyclonal antibody preparations are ones that contain only antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred immunogen compositions are those that contain no other human proteins such as, for example, immunogen compositions made using a non-human host cell for recombinant expression

of a polypeptide of the invention. In such a manner, the only human epitope or epitopes recognized by the resulting antibody compositions raised against this immunogen will be present as part of a polypeptide or polypeptides of the invention.

The antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized polypeptide. If desired, the antibody molecules can be isolated from the mammal (*e.g.*, from the blood) and further purified by well-known techniques, such as protein A chromatography to obtain the IgG fraction. Alternatively, antibodies specific for a protein or polypeptide of the invention can be selected for (*e.g.*, partially purified) or purified by, *e.g.*, affinity chromatography. For example, a recombinantly expressed and purified (or partially purified) protein of the invention is produced as described herein, and covalently or non-covalently coupled to a solid support such as, for example, a chromatography column. The column can then be used to affinity purify antibodies specific for the proteins of the invention from a sample containing antibodies directed against a large number of different epitopes, thereby generating a substantially purified antibody composition, *i.e.*, one that is substantially free of contaminating antibodies. By a substantially purified antibody composition is meant, in this context, that the antibody sample contains at most only 30% (by dry weight) of contaminating antibodies directed against epitopes other than those on the desired protein or polypeptide of the invention, and preferably at most 20%, yet more preferably at most 10%, and most preferably at most 5% (by dry weight) of the sample is contaminating antibodies. A purified antibody composition means that at least 99% of the antibodies in the composition are directed against the desired protein or polypeptide of the invention.

At an appropriate time after immunization, *e.g.*, when the specific antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein, 1975, *Nature* 256:495-497, the human B cell hybridoma technique (Kozbor et al., 1983, *Immunol. Today* 4:72), the EBV-hybridoma technique (Cole et al., 1985, *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing hybridomas is well known (see generally *Current Protocols in Immunology* (1994) Coligan et al. (eds.) John Wiley & Sons, Inc., New York, NY). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the

hybridoma culture supernatants for antibodies that bind the polypeptide of interest, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal antibody directed against a polypeptide of the invention can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with the polypeptide of interest. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia Recombinant Phage Antibody System, Catalog No. 27-9400-01; and the Stratagene SurfZAP Phage Display Kit, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs et al., 1991, *Bio/Technology* 9:1370-1372; Hay et al., 1992, *Hum. Antibod. Hybridomas* 3:81-85; Huse et al., 1989, *Science* 246:1275-1281; Griffiths et al., 1993, *EMBO J.* 12:725-734.

Additionally, recombinant antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. (See, *e.g.*, Cabilly et al., U.S. Patent No. 4,816,567; and Boss et al., U.S. Patent No. 4,816,397, which are incorporated herein by reference in their entirety.) Humanized antibodies are antibody molecules from non-human species having one or more complementarily determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, *e.g.*, Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.) Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better et al., 1988, *Science* 240:1041-1043; Liu et al., 1987, *Proc. Natl. Acad.*

Sci. USA 84:3439-3443; Liu et al., 1987, J. Immunol. 139:3521-3526; Sun et al., 1987, Proc. Natl. Acad. Sci. USA 84:214-218; Nishimura et al., 1987, Canc. Res. 47:999-1005; Wood et al., 1985, Nature 314:446-449; and Shaw et al., 1988, J. Natl. Cancer Inst. 80:1553-1559); Morrison, 1985, Science 229:1202-1207; Oi et al., 1986, Bio/Techniques 4:214; U.S. Patent 5,225,539; Jones et al., 1986, Nature 321:552-525; Verhoeyan et al., 1988, Science 239:1534; and Beidler et al., 1988, J. Immunol. 141:4053-4060.

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced, for example, using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. The transgenic mice are immunized in the normal fashion with a selected antigen, e.g., all or a portion of a polypeptide of the invention. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar, 1995, Int. Rev. Immunol. 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, e.g., a murine antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al., 1994, Bio/technology 12:899-903).

Antibody fragments that contain the idiotypes of the complex can be generated by techniques known in the art. For example, such fragments include, but are not limited to, the F(ab')₂ fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragment that can be generated by reducing the disulfide bridges of

the F(ab')₂ fragment; the Fab fragment that can be generated by treating the antibody molecular with papain and a reducing agent; and Fv fragments.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, e.g., ELISA (enzyme-linked immunosorbent assay). To select antibodies specific to a particular domain of the complex, or a derivative thereof, one may assay generated hybridomas for a product that binds to the fragment of the complex, or a derivative thereof, that contains such a domain. For selection of an antibody that specifically binds a complex of the present, or a derivative, or homolog thereof, but which does not specifically bind to the individual proteins of the complex, or a derivative, or homolog thereof, one can select on the basis of positive binding to the complex and a lack of binding to the individual protein components.

Antibodies specific to a domain of the complex, or a derivative, or homolog thereof, are also provided.

The foregoing antibodies can be used in methods known in the art relating to the localization and/or quantification of the complexes of the invention, e.g., for imaging these proteins, measuring levels thereof in appropriate physiological samples (by immunoassay), in diagnostic methods, etc. This hold true also for a derivative, or homolog thereof of a complex.

In another embodiment of the invention (see *infra*), an antibody to a complex or a fragment of such antibodies containing the antibody binding domain, is a Therapeutic.

5.3. DIAGNOSTIC, PROGNOSTIC, AND SCREENING USES OF PROTEIN COMPLEXES

The particular protein complexes of the present invention may be markers of normal physiological processes, and thus have diagnostic utility. Further, definition of particular groups of patients with elevations or deficiencies of a protein complex of the present invention, or wherein the protein complex has a change in protein component composition, can lead to new nosological classifications of diseases, furthering diagnostic ability.

Examples for diseases or disorders in which the complexes provided herein are involved and/or associated with are infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic

leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis and cancer.

Detecting levels of protein complexes, or individual component proteins that form the complexes, or detecting levels of the mRNAs encoding the components of the complex, may be used in diagnosis, prognosis, and/or staging to follow the course of a disease state, to follow a therapeutic response, etc.

A protein complex of the present invention and the individual components of the complex and a derivative, analog or subsequence thereof, encoding nucleic acids (and sequences complementary thereto), and anti-complex antibodies and antibodies directed against individual components that can form the complex, are useful in diagnostics. The foregoing molecules can be used in assays, such as immunoassays, to detect, prognose, diagnose, or monitor various conditions, diseases, and disorders characterized by aberrant levels of a complex or aberrant component composition of a complex, or monitor the treatment of such various conditions, diseases, and disorders.

In particular, such an immunoassay is carried out by a method comprising contacting a sample derived from a patient with an anti-complex antibody under conditions such that immunospecific binding can occur, and detecting or measuring the amount of any immunospecific binding by the antibody. In a specific aspect, such binding of antibody, in tissue sections, can be used to detect aberrant complex localization, or aberrant (e.g., high, low or absent) levels of a protein complex or complexes. In a specific embodiment, an antibody to the complex can be used to assay a patient tissue or serum sample for the presence of the complex, where an aberrant level of the complex is an indication of a diseased condition. By "aberrant levels" is meant increased or decreased levels relative to that present, or a standard level representing that present, in an analogous sample from a portion or fluid of the body, or from a subject not having the disorder.

The immunoassays which can be used include but are not limited to competitive and non-competitive assay systems using techniques such as Western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays, protein A immunoassays, to name but a few known in the art.

Nucleic acids encoding the components of the protein complex and related nucleic acid sequences and subsequences, including complementary sequences, can be used in hybridization assays. The nucleic acid sequences, or subsequences thereof, comprising about at least 8 nucleotides, can be used as hybridization probes. Hybridization assays can be used to detect, prognose, diagnose, or monitor conditions, disorders, or disease states associated with aberrant levels of the mRNAs encoding the components of a complex as described, *supra*. In particular, such a hybridization assay is carried out by a method comprising contacting a sample containing nucleic acid with a nucleic acid probe capable of hybridizing to component protein coding DNA or RNA, under conditions such that hybridization can occur, and detecting or measuring any resulting hybridization.

In specific embodiments, diseases and disorders involving or characterized by aberrant levels of a protein complex or aberrant complex composition can be diagnosed, or its suspected presence can be screened for, or a predisposition to develop such disorders can be detected, by determining the component protein composition of the complex, or detecting aberrant levels of a member of the complex or un-complexed component proteins or encoding nucleic acids, or functional activity including, but not restricted to, binding to an interacting partner, or by detecting mutations in component protein RNA, DNA or protein (*e.g.*, mutations such as translocations, truncations, changes in nucleotide or amino acid sequence relative to wild-type that cause increased or decreased expression or activity of a complex, and/or component protein. Such diseases and disorders include, but are not limited to, those described in Section 5.4 and its subsections.

By way of example, levels of a protein complex and the individual components of a complex can be detected by immunoassay, levels of component protein RNA or DNA can be detected by hybridization assays (*e.g.*, Northern blots, dot blots, RNase protection assays), and binding of component proteins to each other (*e.g.*, complex formation) can be measured by binding assays commonly known in the art. Translocations and point mutations in component protein genes can be detected by Southern blotting, RFLP analysis, PCR using primers that preferably generate a fragment spanning at least most of the gene by sequencing of genomic DNA or cDNA obtained from the patient, etc.

Assays well known in the art (*e.g.*, assays described above such as immunoassays, nucleic acid hybridization assays, activity assays, etc.) can be used to

determine whether one or more particular protein complexes are present at either increased or decreased levels, or are absent, in samples from patients suffering from a particular disease or disorder, or having a predisposition to develop such a disease or disorder, as compared to the levels in samples from subjects not having such a disease or disorder, or having a predisposition to develop such a disease or disorder.

Additionally, these assays can be used to determine whether the ratio of the complex to the un-complexed components of the complex, is increased or decreased in samples from patients suffering from a particular disease or disorder, or having a predisposition to develop such a disease or disorder, as compared to the ratio in samples from subjects not having such a disease or disorder. In the event that levels of one or more particular protein complexes (*i.e.*, complexes formed from component protein derivatives, homologs, fragments, or analogs) are determined to be increased in patients suffering from a particular disease or disorder, or having a predisposition to develop such a disease or disorder, then the particular disease or disorder, or predisposition for a disease or disorder, can be diagnosed, have prognosis defined for, be screened for, or be monitored by detecting increased levels of the one or more protein complexes, increased levels of the mRNA that encodes one or more members of the one or more particular protein complexes, or by detecting increased complex functional activity.

Accordingly, in a specific embodiment of the present invention, diseases and disorders involving increased levels of one or more protein complexes can be diagnosed, or their suspected presence can be screened for, or a predisposition to develop such disorders can be detected, by detecting increased levels of the one or more protein complexes, the mRNA encoding both members of the complex, or complex functional activity, or by detecting mutations in the component proteins that stabilize or enhance complex formation, *e.g.*, mutations such as translocations in nucleic acids, truncations in the gene or protein, changes in nucleotide or amino acid sequence relative to wild-type, that stabilize or enhance complex formation.

In the event that levels of one or more particular protein complexes are determined to be decreased in patients suffering from a particular disease or disorder, or having a predisposition to develop such a disease or disorder, then the particular disease or disorder or predisposition for a disease or disorder can be diagnosed, have its prognosis determined, be screened for, or be monitored by detecting decreased levels of the one or more protein complexes, the mRNA that encodes one or more members of

the particular one or more protein complexes, or by detecting decreased protein complex functional activity.

Accordingly, in a specific embodiment of the invention, diseases and disorders involving decreased levels of one or more protein complexes can be diagnosed, or their suspected presence can be screened for, or a predisposition to develop such disorders can be detected, by detecting decreased levels of the one or more protein complexes, the mRNA encoding one or more members of the one or more complexes, or complex functional activity, or by detecting mutations in the component proteins that decrease complex formation, e.g., mutations such as translocations in nucleic acids, truncations in the gene or protein, changes in nucleotide or amino acid sequence relative to wild-type, that decrease complex formation.

Accordingly, in a specific embodiment of the invention, diseases and disorders involving aberrant compositions of the complexes can be diagnosed, or their suspected presence can be screened for, or a predisposition to develop such disorders can be detected, by detecting the component proteins of one or more complexes, or the mRNA encoding the members of the one or more complexes.

The use of detection techniques, especially those involving antibodies against a protein complex, provides a method of detecting specific cells that express the complex or component proteins. Using such assays, specific cell types can be defined in which one or more particular protein complexes are expressed, and the presence of the complex or component proteins can be correlated with cell viability, state, health, etc.

Also embodied are methods to detect a protein complex of the present invention in cell culture models that express particular protein complexes or derivatives thereof, for the purpose of characterizing or preparing the complexes for harvest. This embodiment includes cell sorting of prokaryotes such as but not restricted to bacteria (Davey and Kell, 1996, *Microbiol. Rev.* 60:641-696), primary cultures and tissue specimens from eukaryotes, including mammalian species such as human (Steele et al., 1996, *Clin. Obstet. Gynecol* 39:801-813), and continuous cell cultures (Orfao and Ruiz-Arguelles, 1996, *Clin. Biochem.* 29:5-9). Such isolations can be used as methods of diagnosis, described, *supra*.

5.4. THERAPEUTIC USES OF PROTEIN COMPLEXES

The present invention is directed to a method for treatment or prevention of various diseases and disorders by administration of a therapeutic compound (termed

herein "Therapeutic"). Such "Therapeutics" include, but are not limited to, a protein complex of the present invention, the individual component proteins, and analogs and derivatives (including fragments) of the foregoing (*e.g.*, as described hereinabove); antibodies thereto (as described hereinabove); nucleic acids encoding the component protein, and analogs or derivatives, thereof (*e.g.*, as described hereinabove); component protein antisense nucleic acids, and agents that modulate complex formation and/or activity (*i.e.*, agonists and antagonists).

The protein complexes, as identified herein, are implicated significantly in normal physiological processes such as RNA processing and modification..

Furthermore, the protein complexes as identified herein are implicated in processes which are implicated in or associated with pathological conditions.

Diseases and disorders which can be treated and/or prevented and/or diagnosed by Therapeutics interacting with any of the complexes provided herein are for example infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis and cancer.

These disorders are treated or prevented by administration of a Therapeutic that modulates (*i.e.* inhibits or promotes) protein complex activity or formation. Diseases or disorders associated with aberrant levels of complex activity or formation, or aberrant levels or activity of the component proteins, or aberrant complex composition, may be treated by administration of a Therapeutic that modulates complex formation or activity or by the administration of a protein complex.

Therapeutic may also be administered to modulate complex formation or activity or level thereof in a microbial organism such as yeast, fungi such as candida albicans causing an infectious disease in animals or humans.

Diseases and disorders characterized by increased (relative to a subject not suffering from the disease or disorder) complex levels or activity can be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) complex formation or activity. Therapeutics that can be used include, but are not limited to, the component proteins or an analog, derivative or fragment of the component protein; anti-complex antibodies (*e.g.*, antibodies specific for the protein complex, or a fragment or derivative of the antibody containing the binding region thereof; nucleic acids encoding the component proteins; antisense nucleic acids complementary to nucleic acids encoding the component proteins; and nucleic acids encoding the component protein that are

dysfunctional due to, *e.g.*, a heterologous insertion within the protein coding sequence, that are used to "knockout" endogenous protein function by homologous recombination, see, *e.g.*, Capecchi, 1989, Science 244:1288-1292. In one embodiment, a Therapeutic is 1, 2 or more antisense nucleic acids which are complementary to 1, 2, or more nucleic acids, respectfully, that encode component proteins of a complex.

In a specific embodiment of the present invention, a nucleic acid containing a portion of a component protein gene in which gene sequences flank (are both 5' and 3' to) a different gene sequence, is used as a component protein antagonist, or to promote component protein inactivation by homologous recombination (see also, Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342: 435-438). Additionally, mutants or derivatives of a component protein that has greater affinity for another component protein or the complex than wild type may be administered to compete with wild type protein for binding, thereby reducing the levels of complexes containing the wild type protein. Other Therapeutics that inhibit complex function can be identified by use of known convenient *in vitro* assays, *e.g.*, based on their ability to inhibit complex formation, or as described in Section 5.5, *infra*.

In specific embodiments, Therapeutics that antagonize complex formation or activity are administered therapeutically, including prophylactically, (1) in diseases or disorders involving an increased (relative to normal or desired) level of a complex, for example, in patients where complexes are overactive or overexpressed; or (2) in diseases or disorders where an *in vitro* (or *in vivo*) assay (see *infra*) indicates the utility of antagonist administration. Increased levels of a complex can be readily detected, *e.g.*, by quantifying protein and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or protein levels, or structure and/or activity of the expressed complex (or the encoding mRNA). Many methods standard in the art can be thus employed including, but not limited to, immunoassays to detect complexes and/or visualize complexes (*e.g.*, Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis [SDS-PAGE], immunocytochemistry, etc.), and/or hybridization assays to detect concurrent expression of component protein mRNA (*e.g.*, Northern assays, dot blot analysis, *in situ* hybridization, etc.).

A more specific embodiment of the present invention is directed to a method of reducing complex expression (*i.e.*, expression of the protein components of the complex and/or formation of the complex) by targeting mRNAs that express the protein moieties.

RNA therapeutics currently fall within three classes, antisense species, ribozymes, or RNA aptamers (Good et al., 1997, Gene Therapy 4:45-54).

Antisense oligonucleotides have been the most widely used. By way of example, but not limitation, antisense oligonucleotide methodology to reduce complex formation is presented below, *infra*. Ribozyme therapy involves the administration, induced expression, etc. of small RNA molecules with enzymatic ability to cleave, bind, or otherwise inactivate specific RNAs, to reduce or eliminate expression of particular proteins (Grassi and Marini, 1996, Annals of Medicine 28:499-510; Gibson, 1996, Cancer and Metastasis Reviews 15:287-299). RNA aptamers are specific RNA ligand proteins, such as for Tat and Rev RNA (Good et al., 1997, Gene Therapy 4:45-54) that can specifically inhibit their translation. Aptamers specific for component proteins can be identified by many methods well known in the art, for example, by affecting the formation of a complex in the protein-protein interaction assay described, *infra*.

In another embodiment, the activity or levels of a component protein are reduced by administration of another component protein, or the encoding nucleic acid, or an antibody that immunospecifically binds to the component protein, or a fragment or a derivative of the antibody containing the binding domain thereof.

In another aspect of the invention, diseases or disorders associated with increased levels of a component protein of the complex may be treated or prevented by administration of a Therapeutic that increases complex formation if the complex formation acts to reduce or inactivate the component protein through complex formation. Such diseases or disorders can be treated or prevented by administration of one component member of the complex, administration of antibodies or other molecules that stabilize the complex, etc.

Diseases and disorders associated with underexpression of a complex, or a component protein, are treated or prevented by administration of a Therapeutic that promotes (*i.e.*, increases or supplies) complex levels and/or function, or individual component protein function. Examples of such a Therapeutic include but are not limited to a complex or a derivative, analog or fragment of the complex that are functionally active (*e.g.*, able to form a complex), un-complexed component proteins and derivatives, analogs, and fragments of un-complexed component proteins, and nucleic acids encoding the members of a complex or functionally active derivatives or fragments of the members of the complex, *e.g.*, for use in gene therapy. In a specific embodiment, a Therapeutic includes derivatives, homologs or fragments of a component protein that

increase and/or stabilize complex formation. Examples of other agonists can be identified using *in vitro* assays or animal models, examples of which are described, *infra*.

In yet other specific embodiments of the present invention, Therapeutics that promote complex function are administered therapeutically, including prophylactically, (1) in diseases or disorders involving an absence or decreased (relative to normal or desired) level of a complex, for example, in patients where a complex, or the individual components necessary to form the complex, is lacking, genetically defective, biologically inactive or underactive, or under-expressed; or (2) in diseases or disorders wherein an *in vitro* or *in vivo* assay (see, *infra*) indicates the utility of complex agonist administration. The absence or decreased level of a complex, component protein or function can be readily detected, e.g., by obtaining a patient tissue sample (e.g., from biopsy tissue) and assaying it *in vitro* for RNA or protein levels, structure and/or activity of the expressed complex and/or the concurrent expression of mRNA encoding the two components of the complex. Many methods standard in the art can be thus employed, including but not limited to immunoassays to detect and/or visualize a complex, or the individual components of a complex (e.g., Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis [SDS-PAGE], immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs encoding the individual protein components of a complex by detecting and/or visualizing component mRNA concurrently or separately using, e.g., Northern assays, dot blot analysis, *in situ* hybridization, etc.

In specific embodiments, the activity or levels of a component protein are increased by administration of another component protein of the same complex, or a derivative, homolog or analog thereof, a nucleic acid encoding the other component, or an agent that stabilizes or enhances the other component, or a fragment or derivative of such an agent.

Generally, administration of products of species origin or species reactivity (in the case of antibodies) that is the same species as that of the patient is preferred. Thus, in a preferred embodiment, a human complex, or derivative, homolog or analog thereof; nucleic acids encoding the members of the human complex or a derivative, homolog or analog thereof; an antibody to a human complex, or a derivative thereof; or other human agents that affect component proteins or the complex, are therapeutically or prophylactically administered to a human patient.

Preferably, suitable *in vitro* or *in vivo* assays are utilized to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue or individual.

In various specific embodiments, *in vitro* assays can be carried out with representative cells of cell types involved in a patient's disorder, to determine if a Therapeutic has a desired effect upon such cell types.

Compounds for use in therapy can be tested in suitable animal model systems prior to testing in humans, including, but not limited to, rats, mice, chicken, cows, monkeys, rabbits, etc. For *in vivo* testing, prior to administration to humans, any animal model system known in the art may be used. Additional descriptions and sources of Therapeutics that can be used according to the invention are found in Sections 5.1 to 5.3 and 5.7 herein.

5.4.1. GENE THERAPY

In a specific embodiment of the present invention, nucleic acids comprising a sequence encoding the component proteins, or a functional derivative thereof, are administered to modulate complex activity or formation by way of gene therapy. Gene therapy refers to therapy performed by the administration of a nucleic acid to a subject. In this embodiment of the present invention, the nucleic acid expresses its encoded protein(s) that mediates a therapeutic effect by modulating complex activity or formation. Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

For general reviews of the methods of gene therapy, see Goldspiel et al., 1993, *Clinical Pharmacy* 12:488-505; Wu and Wu, 1991, *Biotherapy* 3:87-95; Tolstoshev, 1993, *Ann. Rev. Pharmacol. Toxicol.* 32:573-596; Mulligan, 1993, *Science* 260:926-932; Morgan and Anderson, 1993, *Ann. Rev. Biochem.* 62:191-217; and May, 1993, *TIBTECH* 11:155-215. Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al., eds., 1993, *Current Protocols in Molecular Biology*, John Wiley & Sons, NY; and Kriegler, 1990, *Gene Transfer and Expression*, A Laboratory Manual, Stockton Press, NY.

In a preferred aspect, the Therapeutic comprises a nucleic acid that is part of an expression vector that expresses one or more of the component proteins, or fragments or chimeric proteins thereof, in a suitable host. In particular, such a nucleic acid has a promoter operably linked to the protein coding region(s) (or, less preferably separate

promoters linked to the separate coding regions separately), said promoter being inducible or constitutive, and optionally, tissue-specific. In another particular embodiment, a nucleic acid molecule is used in which the coding sequences, and any other desired sequences, are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intra-chromosomal expression of the component protein nucleic acids (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438).

Delivery of the nucleic acid into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector, or indirect, in which case, cells are first transformed with the nucleic acid *in vitro*, then transplanted into the patient. These two approaches are known, respectively, as *in vivo* or *ex vivo* gene therapy.

In a specific embodiment, the nucleic acid is directly administered *in vivo*, where it is expressed to produce the encoded product. This can be accomplished by any of numerous methods known in the art, e.g., by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by infection using a defective or attenuated retroviral or other viral vector (U.S. Patent No. 4,980,286), or by direct injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors, or through use of transfecting agents, by encapsulation in liposomes, microparticles, or microcapsules, or by administering it in linkage to a peptide that is known to enter the nucleus, or by administering it in linkage to a ligand subject to receptor-mediated endocytosis that can be used to target cell types specifically expressing the receptors (e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432), etc. In another embodiment, a nucleic acid-ligand complex can be formed in which the ligand comprises a fusogenic viral peptide that disrupts endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted *in vivo* for cell specific uptake and expression, by targeting a specific receptor (see, e.g., International Patent Publications WO 92/06180; WO 92/22635; WO 92/20316; WO 93/14188; and WO 93/20221. Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438).

In a specific embodiment, a viral vector that contains the component protein encoding nucleic acids is used. For example, a retroviral vector can be used (Miller et al., 1993, *Meth. Enzymol.* 217:581-599). These retroviral vectors have been modified to delete retroviral sequences that are not necessary for packaging of the viral genome and integration into host cell DNA. The encoding nucleic acids to be used in gene therapy is/are cloned into the vector, which facilitates delivery of the gene into a patient. More detail about retroviral vectors can be found in Boesen et al., 1994, *Biotherapy* 6:291-302, which describes the use of a retroviral vector to deliver the *mdr1* gene to hematopoietic stem cells in order to make the stem cells more resistant to chemotherapy. Other references illustrating the use of retroviral vectors in gene therapy are Clowes et al., 1994, *J. Clin. Invest.* 93:644-651; Kiem et al., 1994, *Blood* 83:1467-1473; Salmons and Gunzberg, 1993, *Human Gene Therapy* 4:129-141; and Grossman and Wilson, 1993, *Curr. Opin. in Genetics and Devel.* 3:110-114.

Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are the liver, the central nervous system, endothelial cells and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Kozarsky and Wilson, 1993, *Current Opinion in Genetics and Development* 3:499-503, discuss adenovirus-based gene therapy. The use of adenovirus vectors to transfer genes to the respiratory epithelia of rhesus monkeys has been demonstrated by Bout et al., 1994, *Human Gene Therapy* 5:3-10. Other instances of the use of adenoviruses in gene therapy can be found in Rosenfeld et al., 1991, *Science* 252:431-434; Rosenfeld et al., 1992, *Cell* 68:143-155; and Mastrangeli et al., 1993, *J. Clin. Invest.* 91:225-234.

Adeno-associated virus (AAV) has also been proposed for use in gene therapy (Walsh et al., 1993, *Proc. Soc. Exp. Biol. Med.* 204:289-300).

Another approach to gene therapy involves transferring a gene into cells in tissue culture by methods such as electroporation, lipofection, calcium phosphate-mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene from those that have not. Those cells are then delivered to a patient.

In this embodiment, the nucleic acid is introduced into a cell prior to administration *in vivo* of the resulting recombinant cell. Such introduction can be carried out by any method known in the art including, but not limited to, transfection by electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells (see, e.g., Loeffler and Behr, 1993, Meth. Enzymol. 217:599-618; Cohen et al., 1993, Meth. Enzymol. 217:618-644; Cline, 1985, Pharmac. Ther. 29:69-92) and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably, is heritable and expressible by its cell progeny.

The resulting recombinant cells can be delivered to a patient by various methods known in the art. In a preferred embodiment, epithelial cells are injected, e.g., subcutaneously. In another embodiment, recombinant skin cells may be applied as a skin graft onto the patient. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be determined by one skilled in the art.

Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes, blood cells such as T lymphocytes, B lymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, and granulocytes, various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, etc.

In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

In an embodiment in which recombinant cells are used in gene therapy, a component protein encoding nucleic acid is/are introduced into the cells such that the gene or genes are expressible by the cells or their progeny, and the recombinant cells are then administered *in vivo* for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem and/or progenitor cells which can be isolated and

maintained *in vitro* can potentially be used in accordance with this embodiment of the present invention. Such stem cells include but are not limited to hematopoietic stem cells (HSCs), stem cells of epithelial tissues such as the skin and the lining of the gut, embryonic heart muscle cells, liver stem cells (International Patent Publication WO 94/08598), and neural stem cells (Stemple and Anderson, 1992, Cell 71:973-985).

Epithelial stem cells (ESCs), or keratinocytes, can be obtained from tissues such as the skin and the lining of the gut by known procedures (Rheinwald, 1980, Meth. Cell Biol. 2A:229). In stratified epithelial tissue such as the skin, renewal occurs by mitosis of stem cells within the germinal layer, the layer closest to the basal lamina. Similarly, stem cells within the lining of the gut provide for a rapid renewal rate of this tissue. ESCs or keratinocytes obtained from the skin or lining of the gut of a patient or donor can be grown in tissue culture (Rheinwald, 1980, Meth. Cell Bio. 2A:229; Pittelkow and Scott, 1986, Mayo Clinic Proc. 61:771). If the ESCs are provided by a donor, a method for suppression of host versus graft reactivity (e.g., irradiation, or drug or antibody administration to promote moderate immunosuppression) can also be used.

With respect to hematopoietic stem cells (HSCs), any technique that provides for the isolation, propagation, and maintenance *in vitro* of HSCs can be used in this embodiment of the invention. Techniques by which this may be accomplished include (a) the isolation and establishment of HSC cultures from bone marrow cells isolated from the future host, or a donor, or (b) the use of previously established long-term HSC cultures, which may be allogeneic or xenogeneic. Non-autologous HSCs are used preferably in conjunction with a method of suppressing transplantation immune reactions between the future host and patient. In a particular embodiment of the present invention, human bone marrow cells can be obtained from the posterior iliac crest by needle aspiration (see, e.g., Kodo et al., 1984, J. Clin. Invest. 73: 1377-1384). In a preferred embodiment of the present invention, the HSCs can be made highly enriched or in substantially pure form. This enrichment can be accomplished before, during, or after long-term culturing, and can be done by any technique known in the art. Long-term cultures of bone marrow cells can be established and maintained by using, for example, modified Dexter cell culture techniques (Dexter et al., 1977, J. Cell Physiol. 91:335) or Witlock-Witte culture techniques (Witlock and Witte, 1982, Proc. Natl. Acad. Sci. USA 79:3608-3612).

In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding region, such that

expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription.

Additional methods can be adapted for use to deliver a nucleic acid encoding the component proteins, or functional derivatives thereof, e.g., as described in Section 5.1, *supra*.

5.4.2. USE OF ANTISENSE OLIGONUCLEOTIDES FOR SUPPRESSION OF PROTEIN COMPLEX ACTIVITY OR FORMATION

In a specific embodiment of the present invention, protein complex activity and formation is inhibited by use of antisense nucleic acids for the component proteins of the complex, that inhibit transcription and/or translation of their complementary sequence. The present invention provides the therapeutic or prophylactic use of nucleic acids of at least six nucleotides that are antisense to a gene or cDNA encoding a component protein, or a portion thereof. An "antisense" nucleic acid as used herein refers to a nucleic acid capable of hybridizing to a sequence-specific portion of a component protein RNA (preferably mRNA) by virtue of some sequence complementarity. The antisense nucleic acid may be complementary to a coding and/or noncoding region of a component protein mRNA. Such antisense nucleic acids that inhibit complex formation or activity have utility as Therapeutics, and can be used in the treatment or prevention of disorders as described *supra*.

The antisense nucleic acids of the invention can be oligonucleotides that are double-stranded or single-stranded, RNA or DNA, or a modification or derivative thereof, which can be directly administered to a cell, or which can be produced intracellularly by transcription of exogenous, introduced sequences.

In another embodiment, the present invention is directed to a method for inhibiting the expression of component protein nucleic acid sequences, in a prokaryotic or eukaryotic cell, comprising providing the cell with an effective amount of a composition comprising an antisense nucleic acid of the component protein, or a derivative thereof, of the invention.

The antisense nucleic acids are of at least six nucleotides and are preferably oligonucleotides, ranging from 6 to about 200 nucleotides. In specific aspects, the oligonucleotide is at least 10 nucleotides, at least 15 nucleotides, at least 100 nucleotides, or at least 200 nucleotides. The oligonucleotides can be DNA or RNA or chimeric mixtures, or derivatives or modified versions thereof, and either single-stranded

or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone. The oligonucleotide may include other appending groups such as peptides, agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; International Patent Publication No. WO 88/09810) or blood-brain barrier (see, e.g., International Patent Publication No. WO 89/10134), hybridization-triggered cleavage agents (see, e.g., Krol et al., 1988, BioTechniques 6:958-976), or intercalating agents (see, e.g., Zon, 1988, Pharm. Res. 5:539-549).

In a preferred aspect of the invention, an antisense oligonucleotide is provided, preferably as single-stranded DNA. The oligonucleotide may be modified at any position in its structure with constituents generally known in the art.

The antisense oligonucleotides may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl)uracil, 5-carboxymethylaminomethyl-2-thio-uridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5N-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methyl-thio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

In another embodiment, the oligonucleotide comprises at least one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal, or an analog of the foregoing.

In yet another embodiment, the oligonucleotide is a 2'-anomeric oligonucleotide. An a-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641).

The oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization-triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

Oligonucleotides of the invention may be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligo-nucleotides may be synthesized by the method of Stein et al. (1988, Nucl. Acids Res. 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

In a specific embodiment, the antisense oligonucleotides comprise catalytic RNAs, or ribozymes (see, e.g., International Patent Publication No. WO 90/11364; Sarver et al., 1990, Science 247:1222-1225). In another embodiment, the oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analog (Inoue et al., 1987, FEBS Lett. 215:327-330).

In an alternative embodiment, the antisense nucleic acids of the invention are produced intracellularly by transcription from an exogenous sequence. For example, a vector can be introduced in vivo such that it is taken up by a cell, within which cell the vector or a portion thereof is transcribed, producing an antisense nucleic acid (RNA) of the invention. Such a vector would contain a sequence encoding the component protein. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art to be capable of replication and expression in mammalian cells. Expression of the sequences encoding the antisense RNAs can be by any promoter known in the art to act in mammalian, preferably human, cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to, the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto

et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al., 1982, Nature 296:39-42), etc.

The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a component protein gene, preferably a human gene. However, absolute complementarity, although preferred, is not required. A sequence "complementary to at least a portion of an RNA," as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with a component protein RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

The component protein antisense nucleic acids can be used to treat (or prevent) disorders of a cell type that expresses, or preferably overexpresses, a protein complex.

Cell types that express or overexpress component protein RNA can be identified by various methods known in the art. Such methods include, but are not limited to, hybridization with component protein-specific nucleic acids (e.g., by Northern blot hybridization, dot blot hybridization, or *in situ* hybridization), or by observing the ability of RNA from the cell type to be translated *in vitro* into the component protein by immunohistochemistry, Western blot analysis, ELISA, etc. In a preferred aspect, primary tissue from a patient can be assayed for protein expression prior to treatment, e.g., by immunocytochemistry, *in situ* hybridization, or any number of methods to detect protein or mRNA expression.

Pharmaceutical compositions of the invention (see Section 5.7, *infra*), comprising an effective amount of a protein component antisense nucleic acid in a pharmaceutically acceptable carrier can be administered to a patient having a disease or disorder that is of a type that expresses or overexpresses a protein complex of the present invention.

The amount of antisense nucleic acid that will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. Where possible, it is desirable to

determine the antisense cytotoxicity in vitro, and then in useful animal model systems, prior to testing and use in humans.

In a specific embodiment, pharmaceutical compositions comprising antisense nucleic acids are administered via liposomes, microparticles, or microcapsules. In various embodiments of the invention, it may be useful to use such compositions to achieve sustained release of the antisense nucleic acids. In a specific embodiment, it may be desirable to utilize liposomes targeted via antibodies to specific identifiable central nervous system cell types (Leonetti et al., 1990, Proc. Natl. Acad. Sci. U.S.A. 87:2448-2451; Renneisen et al., 1990, J. Biol. Chem. 265:16337-16342).

5.5. ASSAYS OF PROTEIN COMPLEXES AND DERIVATIVES AND ANALOGS THEREOF

The functional activity of a protein complex of the present invention, or a derivative, fragment or analog thereof, can be assayed by various methods. Potential modulators (e.g., agonists and antagonists) of complex activity or formation, e.g., anti-complex antibodies and antisense nucleic acids, can be assayed for the ability to modulate complex activity or formation.

In one embodiment of the present invention, where one is assaying for the ability to bind or compete with a wild-type complex for binding to an anti-complex antibody, various immunoassays known in the art can be used, including but not limited to competitive and non-competitive assay systems using techniques such as radioimmunoassay, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, in situ immunoassays (using colloidal gold, enzyme or radioisotope labels), western blot analysis, precipitation reactions, agglutination assays (e.g., gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention.

The expression of the component protein genes (both endogenous and those expressed from cloned DNA containing the genes) can be detected using techniques

known in the art, including but not limited to Southern hybridization (Southern, 1975, J. Mol. Biol. 98:503-517), northern hybridization (see, e.g., Freeman et al., 1983, Proc. Natl. Acad. Sci. USA 80:4094-4098), restriction endonuclease mapping (Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press, New York), RNase protection assays (Current Protocols in Molecular Biology, John Wiley and Sons, New York, 1997), DNA sequence analysis, and polymerase chain reaction amplification (PCR; U.S. Patent Nos. 4,683,202, 4,683,195, and 4,889,818; Gyllenstein et al., 1988, Proc. Natl. Acad. Sci. USA 85:7652-7657; Ochman et al., 1988, Genetics 120:621-623; Loh et al., 1989, Science 243:217-220) followed by Southern hybridization with probes specific for the component protein genes, in various cell types. Methods of amplification other than PCR commonly known in the art can be employed. In one embodiment, Southern hybridization can be used to detect genetic linkage of component protein gene mutations to physiological or pathological states. Various cell types, at various stages of development, can be characterized for their expression of component proteins at the same time and in the same cells. The stringency of the hybridization conditions for northern or Southern blot analysis can be manipulated to ensure detection of nucleic acids with the desired degree of relatedness to the specific probes used. Modifications to these methods and other methods commonly known in the art can be used.

Derivatives (e.g., fragments), homologs and analogs of one component protein can be assayed for binding to another component protein in the same complex by any method known in the art, for example the modified yeast matrix mating test described in Section 5.6.1 *infra*, immunoprecipitation with an antibody that binds to the component protein complexed with other component proteins in the same complex, followed by size fractionation of the immunoprecipitated proteins (e.g., by denaturing or nondenaturing polyacrylamide gel electrophoresis), Western blot analysis, etc.

One embodiment of the invention provides a method for screening a derivative, homolog or analog of a component protein for biological activity comprising contacting said derivative, homolog or analog of the component protein with the other component proteins in the same complex; and detecting the formation of a complex between said derivative, homolog or analog of the component protein and the other component proteins; wherein detecting formation of said complex indicates that said derivative, homolog or analog of has biological (e.g., binding) activity.

The invention also provides methods of modulating the activity of a component protein that can participate in a protein complex by administration of a binding partner of that protein or derivative, homolog or analog thereof.

In a specific embodiment of the present invention, a protein complex of the present invention is administered to treat or prevent a disease or disorder, since the complex and/or component proteins have been implicated in the disease and disorder. Accordingly, a protein complex or a derivative, homolog, analog or fragment thereof, nucleic acids encoding the component proteins, anti-complex antibodies, and other modulators of protein complex activity, can be tested for activity in treating or preventing a disease or disorder in *in vitro* and *in vivo* assays.

In one embodiment, a Therapeutic of the invention can be assayed for activity in treating or preventing a disease by contacting cultured cells that exhibit an indicator of the disease *in vitro*, with the Therapeutic, and comparing the level of said indicator in the cells contacted with the Therapeutic, with said level of said indicator in cells not so contacted, wherein a lower level in said contacted cells indicates that the Therapeutic has activity in treating or preventing the disease.

In another embodiment of the invention, a Therapeutic of the invention can be assayed for activity in treating or preventing a disease by administering the Therapeutic to a test animal that is predisposed to develop symptoms of a disease, and measuring the change in said symptoms of the disease after administration of said Therapeutic, wherein a reduction in the severity of the symptoms of the disease or prevention of the symptoms of the disease indicates that the Therapeutic has activity in treating or preventing the disease. Such a test animal can be any one of a number of animal models known in the art for disease. These animal models are well known in the art. These animal models include, but are not limited to those which are listed in the section 5.6 (*supra*) as exemplary animal models to study any of the complexes provided in the invention.

5.6 SCREENING FOR MODULATORS OF THE PROTEIN COMPLEXES

A complex of the present invention, the component proteins of the complex and nucleic acids encoding the component proteins, as well as derivatives and fragments of the amino and nucleic acids, can be used to screen for compounds that bind to, or modulate the amount of, activity of, or protein component composition of, said complex, and thus, have potential use as modulators, *i.e.*, agonists or antagonists, of complex

activity, and/or complex formation, *i.e.*, the amount of complex formed, and/or protein component composition of the complex.

Thus, the present invention is also directed to methods for screening for molecules that bind to, or modulate the amount of, activity of, or protein component composition of, a complex of the present invention. In one embodiment of the invention, the method for screening for a molecule that modulates directly or indirectly the function, activity or formation of a complex of the present invention comprises exposing said complex, or a cell or organism containing the complex machinery, to one or more candidate molecules under conditions conducive to modulation; and determining the amount of, activity of, or identities of the protein components of, said complex, wherein a change in said amount, activity, or identities relative to said amount, activity or identities in the absence of said candidate molecules indicates that the molecules modulate function, activity or formation of said complex.

In another embodiment, the present invention further relates to a process for the identification and/or preparation of an effector of the cleavage/polyadenylation of precursor mRNA comprising the step of bringing into contact a product of any of claims 1 to 7 with a compound, a mixture or a library of compounds and determining whether the compound or a certain compound of the mixture or library binds to the product and/or effects the products biological activity and optionally further purifying the compound positively tested as effector.

In another embodiment, the present invention is directed to a method for screening for a molecule that binds a protein complex of the present invention comprising exposing said complex, or a cell or organism containing the complex machinery, to one or more candidate molecules; and determining whether said complex is bound by any of said candidate molecules. Such screening assays can be carried out using cell-free and cell-based methods that are commonly known in the art *in vitro*, *in vivo* or *ex vivo*. For example, an isolated complex can be employed, or a cell can be contacted with the candidate molecule and the complex can be isolated from such contacted cells and the isolated complex can be assayed for activity or component composition. In another example, a cell containing the complex can be contacted with the candidate molecule and the levels of the complex in the contacted cell can be measured. Additionally, such assays can be carried out in cells recombinantly expressing a component protein from column A of table 1 of a given row, or a functionally active fragment or functionally active derivative thereof, and a component

protein from column B of table 1 of said row, or a functionally active fragment or functionally active derivative thereof. Additionally, such assays can also be carried out in cells recombinantly expressing all component proteins from the group of proteins in column C of table 1.

For example, assays can be carried out using recombinant cells expressing the protein components of a complex, to screen for molecules that bind to, or interfere with, or promote complex activity or formation. In preferred embodiments, polypeptide derivatives that have superior stabilities but retain the ability to form a complex (*e.g.*, one or more component proteins modified to be resistant to proteolytic degradation in the binding assay buffers, or to be resistant to oxidative degradation), are used to screen for modulators of complex activity or formation. Such resistant molecules can be generated, *e.g.*, by substitution of amino acids at proteolytic cleavage sites, the use of chemically derivatized amino acids at proteolytic susceptible sites, and the replacement of amino acid residues subject to oxidation, *i.e.* methionine and cysteine.

A particular aspect of the present invention relates to identifying molecules that inhibit or promote formation or degradation of a complex of the present invention, *e.g.*, using the method described for isolating the complex and identifying members of the complex using the TAP assay described in Section 6, *infra*, and in WO 00/09716 and Rigaut et al., 1999, Nature Biotechnology 17:1030-1032, which are each incorporated by reference in their entirety.

In another embodiment of the invention, a modulator is identified by administering a candidate molecule to a transgenic non-human animal expressing the complex component proteins from promoters that are not the native promoters of the respective proteins, more preferably where the candidate molecule is also recombinantly expressed in the transgenic non-human animal. Alternatively, the method for identifying such a modulator can be carried out *in vitro*, preferably with a purified complex, and a purified candidate molecule.

Agents/molecules (candidate molecules) to be screened can be provided as mixtures of a limited number of specified compounds, or as compound libraries, peptide libraries and the like. Agents/molecules to be screened may also include all forms of antisera, antisense nucleic acids, etc., that can modulate complex activity or formation. Exemplary candidate molecules and libraries for screening are set forth in Section 5.6.1, *infra*.

Screening the libraries can be accomplished by any of a variety of commonly known methods. See, e.g., the following references, which disclose screening of peptide libraries: Parmley and Smith, 1989, *Adv. Exp. Med. Biol.* 251:215-218; Scott and Smith, 1990, *Science* 249:386-390; Fowlkes et al., 1992, *BioTechniques* 13:422-427; Oldenburg et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:5393-5397; Yu et al., 1994, *Cell* 76:933-945; Staudt et al., 1988, *Science* 241:577-580; Bock et al., 1992, *Nature* 355:564-566; Tuerk et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:6988-6992; Ellington et al., 1992, *Nature* 355:850-852; U.S. Patent No. 5,096,815, U.S. Patent No. 5,223,409, and U.S. Patent No. 5,198,346, all to Ladner et al.; Rebar and Pabo, 1993, *Science* 263:671-673; and International Patent Publication No. WO 94/18318.

In a specific embodiment, screening can be carried out by contacting the library members with a complex immobilized on a solid phase, and harvesting those library members that bind to the protein (or encoding nucleic acid or derivative). Examples of such screening methods, termed "panning" techniques, are described by way of example in Parmley and Smith, 1988, *Gene* 73:305-318; Fowlkes et al., 1992, *BioTechniques* 13:422-427; International Patent Publication No. WO 94/18318; and in references cited hereinabove.

In a specific embodiment, fragments and/or analogs of protein components of a complex, especially peptidomimetics, are screened for activity as competitive or non-competitive inhibitors of complex formation (amount of complex or composition of complex) or activity in the cell, which thereby inhibit complex activity or formation in the cell.

In one embodiment, agents that modulate (*i.e.*, antagonize or agonize) complex activity or formation can be screened for using a binding inhibition assay, wherein agents are screened for their ability to modulate formation of a complex under aqueous, or physiological, binding conditions in which complex formation occurs in the absence of the agent to be tested. Agents that interfere with the formation of complexes of the invention are identified as antagonists of complex formation. Agents that promote the formation of complexes are identified as agonists of complex formation. Agents that completely block the formation of complexes are identified as inhibitors of complex formation.

Methods for screening may involve labeling the component proteins of the complex with radioligands (e.g., ^{125}I or ^3H), magnetic ligands (e.g., paramagnetic beads covalently attached to photobiotin acetate), fluorescent ligands (e.g., fluorescein or rhodamine), or enzyme ligands (e.g., luciferase or beta-galactosidase). The reactants

that bind in solution can then be isolated by one of many techniques known in the art, including but not restricted to, co-immunoprecipitation of the labeled complex moiety using antisera against the unlabeled binding partner (or labeled binding partner with a distinguishable marker from that used on the second labeled complex moiety), immunoaffinity chromatography, size exclusion chromatography, and gradient density centrifugation. In a preferred embodiment, the labeled binding partner is a small fragment or peptidomimetic that is not retained by a commercially available filter. Upon binding, the labeled species is then unable to pass through the filter, providing for a simple assay of complex formation.

Methods commonly known in the art are used to label at least one of the component members of the complex. Suitable labeling methods include, but are not limited to, radiolabeling by incorporation of radiolabeled amino acids, e.g., ^3H -leucine or ^{35}S -methionine, radiolabeling by post-translational iodination with ^{125}I or ^{131}I using the chloramine T method, Bolton-Hunter reagents, etc., or labeling with ^{32}P using phosphorylase and inorganic radiolabeled phosphorous, biotin labeling with photobiotin-acetate and sunlamp exposure, etc. In cases where one of the members of the complex is immobilized, e.g., as described *infra*, the free species is labeled. Where neither of the interacting species is immobilized, each can be labeled with a distinguishable marker such that isolation of both moieties can be followed to provide for more accurate quantification, and to distinguish the formation of homomeric from heteromeric complexes. Methods that utilize accessory proteins that bind to one of the modified interactants to improve the sensitivity of detection, increase the stability of the complex, etc., are provided.

Typical binding conditions are, for example, but not by way of limitation, in an aqueous salt solution of 10-250 mM NaCl, 5-50 mM Tris-HCl, pH 5-8, and 0.5% Triton X-100 or other detergent that improves specificity of interaction. Metal chelators and/or divalent cations may be added to improve binding and/or reduce proteolysis. Reaction temperatures may include 4, 10, 15, 22, 25, 35, or 42 degrees Celsius, and time of incubation is typically at least 15 seconds, but longer times are preferred to allow binding equilibrium to occur. Particular complexes can be assayed using routine protein binding assays to determine optimal binding conditions for reproducible binding.

The physical parameters of complex formation can be analyzed by quantification of complex formation using assay methods specific for the label used, e.g., liquid scintillation counting for radioactivity detection, enzyme activity for enzyme-labeled

moieties, etc. The reaction results are then analyzed utilizing Scatchard analysis, Hill analysis, and other methods commonly known in the arts (see, e.g., *Proteins, Structures, and Molecular Principles*, 2nd Edition (1993) Creighton, Ed., W.H. Freeman and Company, New York).

In a second common approach to binding assays, one of the binding species is immobilized on a filter, in a microtiter plate well, in a test tube, to a chromatography matrix, etc., either covalently or non-covalently. Proteins can be covalently immobilized using any method well known in the art, for example, but not limited to the method of Kadonaga and Tjian, 1986, *Proc. Natl. Acad. Sci. USA* 83:5889-5893, *i.e.*, linkage to a cyanogen-bromide derivatized substrate such as CNBr-Sepharose 4B (Pharmacia). Where needed, the use of spacers can reduce steric hindrance by the substrate. Non-covalent attachment of proteins to a substrate include, but are not limited to, attachment of a protein to a charged surface, binding with specific antibodies, binding to a third unrelated interacting protein, etc.

Assays of agents (including cell extracts or a library pool) for competition for binding of one member of a complex (or derivatives thereof) with another member of the complex labeled by any means (e.g., those means described above) are provided to screen for competitors or enhancers of complex formation.

In specific embodiments, blocking agents to inhibit non-specific binding of reagents to other protein components, or absorptive losses of reagents to plastics, immobilization matrices, etc., are included in the assay mixture. Blocking agents include, but are not restricted to bovine serum albumin, beta-casein, nonfat dried milk, Denhardt's reagent, Ficoll, polyvinylpyrrolidone, nonionic detergents (NP40, Triton X-100, Tween 20, Tween 80, etc.), ionic detergents (e.g., SDS, LDS, etc.), polyethylene glycol, etc. Appropriate blocking agent concentrations allow complex formation.

After binding is performed, unbound, labeled protein is removed in the supernatant, and the immobilized protein retaining any bound, labeled protein is washed extensively. The amount of bound label is then quantified using standard methods in the art to detect the label as described, *supra*.

Moreover, a number of polyadenylation assays are described in the prior art. Such assays can be found in Bienroth, S.E.; Wahle, C.; Suter-Crazzolara, C. and Keller, W. (1991), *J. Biol. Chem.* 266, 19768-19776; Edwards-Gilbert, G. and Milcarek, C. (1995), *Mol. Cell. Biol.* 15, 6420-6429; Wahle, E. (1991), *Cell* 66, 759-768; Christofori, G. and Keller, W. (*Cell*) 54, 875-889.

Exemplary assays useful to measure the 3' end processing activity for mRNA of complex 162 include, but are not limited to those described in Kessler MM et al, 1996, J Biol. Chem. 271: 27167-75, and Butler, S. J. and Platt, T. (1988), Science 242, 1270-1274, and Moore, C.L. and Sharp, P.A. (1985), Cell 41, 845-855

Exemplary assays useful to measure the cleavage step in 3' end processing activity of mRNA of complex 162 include, but are not limited to those described in Ruegsegger U et al., 1996, J Biol Chem 271: 6107-6113.

An exemplary RNA binding assay can be carried out by contacting a complex having RNA binding activity with a radioactive [^{32}P] end-labeled RNA substrate, e.g. a poly (A) RNA, under appropriate conditions and detecting bound protein. The detection of bound protein can be carried out, e.g., by filtrating the solution through a nitrocellulose filter and determining the radioactivity bound to the filter. This assay is based on the retention of nucleic acid-protein complexes on Nitrocellulose whereas free nucleic acid can pass through the filter

(see e.g. Wahle, E., 1991, Methods 66: 759-68)

An exemplary RNA exonuclease assay can be carried out by contacting a complex having RNA exonuclease activity with a radioactivity [32 phosphate] end-labeled RNA substrate under appropriate conditions and detecting the release of free radioactive nucleotides. The detection of free radioactive nucleotides can be carried out, e.g., by adding 20% trichloroacetic acid, filtrating the solution through a filter and measuring the amount of acid-soluble radioactivity

(see e.g. Ross, J., 1999, Methods 17: 52-9)

An exemplary mRNA splicing assay can be carried out by contacting a complex having mRNA splicing activity with a radioactively labeled RNA substrate under appropriate conditions and detecting the release of spliced RNA species. The detection of spliced RNA species can be carried out, e.g., by fractionation of processed RNAs in a glycerol gradient and subsequent analysis by denaturing polyacrylamide gel electrophoresis and visualization by autoradiography.

(see e.g. Schwer, B. and Gross, CH., 1998, Methods 17: 2086-94)

An exemplary rRNA processing assay can be carried out by contacting a complex having rRNA processing activity with an pre-rRNA substrate under appropriate conditions and detecting the release of free processed rRNA species. The detection of processed rRNA species can be carried out, e.g., using a primer extension or northern blotting assay by measuring the size of the rRNA species.

(see e.g. Kressler, D. et al, 1997, Methods 17: 7283-94)

5.6.1. CANDIDATE MOLECULES

Any molecule known in the art can be tested for its ability to modulate (increase or decrease) the amount of, activity of, or protein component composition of a complex of the present invention as detected by a change in the amount of, activity of, or protein component composition of, said complex. By way of example, a change in the amount of the complex can be detected by detecting a change in the amount of the complex that can be isolated from a cell expressing the complex machinery. For identifying a molecule that modulates complex activity, candidate molecules can be directly provided to a cell expressing the complex machinery, or, in the case of candidate proteins, can be provided by providing their encoding nucleic acids under conditions in which the nucleic acids are recombinantly expressed to produce the candidate proteins within the cell expressing the complex machinery, the complex is then isolated from the cell and the isolated complex is assayed for activity using methods well known in the art, not limited to those described, *supra*.

This embodiment of the invention is well suited to screen chemical libraries for molecules which modulate, e.g., inhibit, antagonize, or agonize, the amount of, activity of, or protein component composition of the complex. The chemical libraries can be peptide libraries, peptidomimetic libraries, chemically synthesized libraries, recombinant, e.g., phage display libraries, and *in vitro* translation-based libraries, other non-peptide synthetic organic libraries, etc.

Exemplary libraries are commercially available from several sources (ArQule, Tripos/PanLabs, ChemDesign, Pharmacopoeia). In some cases, these chemical libraries are generated using combinatorial strategies that encode the identity of each member of the library on a substrate to which the member compound is attached, thus allowing direct and immediate identification of a molecule that is an effective modulator. Thus, in many combinatorial approaches, the position on a plate of a compound specifies

that compound's composition. Also, in one example, a single plate position may have from 1-20 chemicals that can be screened by administration to a well containing the interactions of interest. Thus, if modulation is detected, smaller and smaller pools of interacting pairs can be assayed for the modulation activity. By such methods, many candidate molecules can be screened.

Many diversity libraries suitable for use are known in the art and can be used to provide compounds to be tested according to the present invention. Alternatively, libraries can be constructed using standard methods. Chemical (synthetic) libraries, recombinant expression libraries, or polysome-based libraries are exemplary types of libraries that can be used.

The libraries can be constrained or semirigid (having some degree of structural rigidity), or linear or nonconstrained. The library can be a cDNA or genomic expression library, random peptide expression library or a chemically synthesized random peptide library, or non-peptide library. Expression libraries are introduced into the cells in which the assay occurs, where the nucleic acids of the library are expressed to produce their encoded proteins.

In one embodiment, peptide libraries that can be used in the present invention may be libraries that are chemically synthesized *in vitro*. Examples of such libraries are given in Houghten et al., 1991, Nature 354:84-86, which describes mixtures of free hexapeptides in which the first and second residues in each peptide were individually and specifically defined; Lam et al., 1991, Nature 354:82-84, which describes a "one bead, one peptide" approach in which a solid phase split synthesis scheme produced a library of peptides in which each bead in the collection had immobilized thereon a single, random sequence of amino acid residues; Medynski, 1994, Bio/Technology 12:709-710, which describes split synthesis and T-bag synthesis methods; and Gallop et al., 1994, J. Medicinal Chemistry 37(9):1233-1251. Simply by way of other examples, a combinatorial library may be prepared for use, according to the methods of Ohlmeyer et al., 1993, Proc. Natl. Acad. Sci. USA 90:10922-10926; Erb et al., 1994, Proc. Natl. Acad. Sci. USA 91:11422-11426; Houghten et al., 1992, Biotechniques 13:412; Jayawickreme et al., 1994, Proc. Natl. Acad. Sci. USA 91:1614-1618; or Salmon et al., 1993, Proc. Natl. Acad. Sci. USA 90:11708-11712. PCT Publication No. WO 93/20242 and Brenner and Lerner, 1992, Proc. Natl. Acad. Sci. USA 89:5381-5383 describe "encoded combinatorial chemical libraries," that contain oligonucleotide identifiers for each chemical polymer library member.

In a preferred embodiment, the library screened is a biological expression library that is a random peptide phage display library, where the random peptides are constrained (e.g., by virtue of having disulfide bonding).

Further, more general, structurally constrained, organic diversity (e.g., nonpeptide) libraries, can also be used. By way of example, a benzodiazepine library (see e.g., Bunin et al., 1994, Proc. Natl. Acad. Sci. USA 91:4708-4712) may be used.

Conformationally constrained libraries that can be used include but are not limited to those containing invariant cysteine residues which, in an oxidizing environment, cross-link by disulfide bonds to form cystines, modified peptides (e.g., incorporating fluorine, metals, isotopic labels, are phosphorylated, etc.), peptides containing one or more non-naturally occurring amino acids, non-peptide structures, and peptides containing a significant fraction of γ -carboxyglutamic acid.

Libraries of non-peptides, e.g., peptide derivatives (for example, that contain one or more non-naturally occurring amino acids) can also be used. One example of these are peptoid libraries (Simon et al., 1992, Proc. Natl. Acad. Sci. USA 89:9367-9371). *Peptoids are polymers of non-natural amino acids that have naturally occurring side chains attached not to the alpha carbon but to the backbone amino nitrogen.* Since peptoids are not easily degraded by human digestive enzymes, they are advantageously more easily adaptable to drug use. Another example of a library that can be used, in which the amide functionalities in peptides have been permethylated to generate a chemically transformed combinatorial library, is described by Ostresh et al., 1994, Proc. Natl. Acad. Sci. USA 91:11138-11142).

The members of the peptide libraries that can be screened according to the invention are not limited to containing the 20 naturally occurring amino acids. In particular, chemically synthesized libraries and polysome based libraries allow the use of amino acids in addition to the 20 naturally occurring amino acids (by their inclusion in the precursor pool of amino acids used in library production). In specific embodiments, the library members contain one or more non-natural or non-classical amino acids or cyclic peptides. Non-classical amino acids include but are not limited to the D-isomers of the common amino acids, γ -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid; -Abu, -Ahx, 6-amino hexanoic acid; Aib, 2-amino isobutyric acid; 3-amino propionic acid; ornithine; norleucine; norvaline, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β -alanine, designer amino acids such as β -methyl amino acids, γ -methyl amino acids, N-methyl amino acids,

fluoro-amino acids and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

In a specific embodiment, fragments and/or analogs of complexes of the invention, or protein components thereof, especially peptidomimetics, are screened for activity as competitive or non-competitive inhibitors of complex activity or formation.

In another embodiment of the present invention, combinatorial chemistry can be used to identify modulators of a the complexes. Combinatorial chemistry is capable of creating libraries containing hundreds of thousands of compounds, many of which may be structurally similar. While high throughput screening programs are capable of screening these vast libraries for affinity for known targets, new approaches have been developed that achieve libraries of smaller dimension but which provide maximum chemical diversity. (See *e.g.*, Matter, 1997, *Journal of Medicinal Chemistry* 40:1219-1229).

One method of combinatorial chemistry, affinity fingerprinting, has previously been used to test a discrete library of small molecules for binding affinities for a defined panel of proteins. The fingerprints obtained by the screen are used to predict the affinity of the individual library members for other proteins or receptors of interest (in the instant invention, the protein complexes of the present invention and protein components thereof.) The fingerprints are compared with fingerprints obtained from other compounds known to react with the protein of interest to predict whether the library compound might similarly react. For example, rather than testing every ligand in a large library for interaction with a complex or protein component, only those ligands having a fingerprint similar to other compounds known to have that activity could be tested. (See, *e.g.*, Kauvar et al., 1995, *Chemistry and Biology* 2:107-118; Kauvar, 1995, *Affinity fingerprinting*, *Pharmaceutical Manufacturing International*. 8:25-28; and Kauvar, *Toxic-Chemical Detection by Pattern Recognition in New Frontiers in Agrochemical Immunoassay*, D. Kurtz, L. Stanker and J.H. Skerritt, Editors, 1995, AOAC: Washington, D.C., 305-312).

Kay et al., 1993, *Gene* 128:59-65 (Kay) discloses a method of constructing peptide libraries that encode peptides of totally random sequence that are longer than those of any prior conventional libraries. The libraries disclosed in Kay encode totally synthetic random peptides of greater than about 20 amino acids in length. Such libraries can be advantageously screened to identify complex modulators. (See also U.S. Patent

No. 5,498,538 dated March 12, 1996; and PCT Publication No. WO 94/18318 dated August 18, 1994).

A comprehensive review of various types of peptide libraries can be found in Gallop et al., 1994, J. Med. Chem. 37:1233-1251.

5.7. PHARMACEUTICAL COMPOSITIONS AND THERAPEUTIC/PROPHYLACTIC ADMINISTRATION

The invention provides methods of treatment (and prophylaxis) by administration to a subject of an effective amount of a Therapeutic of the invention. In a preferred aspect, the Therapeutic is substantially purified. The subject is preferably an animal including, but not limited to animals such as cows, pigs, horses, chickens, cats, dogs, etc., and is preferably a mammal, and most preferably human. In a specific embodiment, a non-human mammal is the subject.

Various delivery systems are known and can be used to administer a Therapeutic of the invention, e.g., encapsulation in liposomes, microparticles, and microcapsules; use of recombinant cells capable of expressing the Therapeutic, use of receptor-mediated endocytosis (e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432); construction of a Therapeutic nucleic acid as part of a retroviral or other vector, etc. Methods of introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The compounds may be administered by any convenient route, for example by infusion, by bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral, rectal and intestinal mucosa, etc.), and may be administered together with other biologically active agents. Administration can be systemic or local. In addition, it may be desirable to introduce the pharmaceutical compositions of the invention into the central nervous system by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

In a specific embodiment, it may be desirable to administer the pharmaceutical compositions of the invention locally to the area in need of treatment. This may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, or by means of an implant, said

implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. In one embodiment, administration can be by direct injection at the site (or former site) of a malignant tumor or neoplastic or pre-neoplastic tissue.

In another embodiment, the Therapeutic can be delivered in a vesicle, in particular a liposome (Langer, 1990, *Science* 249:1527-1533; Treat et al., 1989, In: *Liposomes in the Therapy of Infectious Disease and Cancer*, Lopez-Berestein and Fidler, eds., Liss, New York, pp. 353-365; Lopez-Berestein, *ibid.*, pp. 317-327; see generally *ibid.*)

In yet another embodiment, the Therapeutic can be delivered via a controlled release system. In one embodiment, a pump may be used (Langer, *supra*; Sefton, 1987, *CRC Crit. Ref. Biomed. Eng.* 14:201-240; Buchwald et al., 1980, *Surgery* 88:507-516; Saudek et al., 1989, *N. Engl. J. Med.* 321:574-579). In another embodiment, polymeric materials can be used (*Medical Applications of Controlled Release*, Langer and Wise, eds., CRC Press, Boca Raton, Florida, 1974; *Controlled Drug Bioavailability, Drug Product Design and Performance*, Smolen and Ball, eds., Wiley, New York, 1984; Ranger and Peppas, 1983, *Macromol. Sci. Rev. Macromol. Chem.* 23:61; Levy et al., 1985, *Science* 228:190-192; During et al., 1989, *Ann. Neurol.* 25:351-356; Howard et al., 1989, *J. Neurosurg.* 71:858-863). In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, *i.e.*, the brain, thus requiring only a fraction of the systemic dose (*e.g.*, Goodson, 1984, In: *Medical Applications of Controlled Release*, *supra*, Vol. 2, pp. 115-138). Other controlled release systems are discussed in the review by Langer (1990, *Science* 249:1527-1533).

In a specific embodiment where the Therapeutic is a nucleic acid encoding a protein Therapeutic, the nucleic acid can be administered *in vivo* to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, *e.g.*, by use of a retroviral vector (U.S. Patent No. 4,980,286), or by direct injection, or by use of microparticle bombardment (*e.g.*, a gene gun; Biolistic, Dupont), or by coating it with lipids, cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (*e.g.*, Joliot et al., 1991, *Proc. Natl. Acad. Sci. USA* 88:1864-1868), etc. Alternatively, a nucleic acid Therapeutic can be introduced intracellularly and incorporated by homologous recombination within host cell DNA for expression.

The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a Therapeutic, and a pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly, in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, including but not limited to peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered orally. Saline and aqueous dextrose are preferred carriers when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions are preferably employed as liquid carriers for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsions, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin. Such compositions will contain a therapeutically effective amount of the Therapeutic, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient. The formulation should suit the mode of administration.

In a preferred embodiment, the composition is formulated, in accordance with routine procedures, as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anesthetic such as lidocaine to ease pain at

the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water-free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water or saline for injection can be provided so that the ingredients may be mixed prior to administration.

The Therapeutics of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with free carboxyl groups such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., those formed with free amine groups such as those derived from isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc., and those derived from sodium, potassium, ammonium, calcium, and ferric hydroxides, etc.

The amount of the Therapeutic of the invention which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. In addition, *in vitro* assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances. However, suitable dosage ranges for intravenous administration are generally about 20-500 micrograms of active compound per kilogram body weight. Suitable dosage ranges for intranasal administration are generally about 0.01 pg/kg body weight to 1 mg/kg body weight. Effective doses may be extrapolated from dose-response curves derived from *in vitro* or animal model test systems.

Suppositories generally contain active ingredient in the range of 0.5% to 10% by weight; oral formulations preferably contain 10% to 95% active ingredient.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. For example, the kit can comprise in one or more containers a first protein, or a functionally active fragment or functionally active derivative thereof, which first protein is selected from the group consisting of proteins listed in column A of table 1 of a given row; and a second protein, or a functionally active fragment or functionally active derivative thereof, which second protein is selected from the group consisting of proteins listed in column B of table 1 of said row. Alternatively, the kit can comprise in one or more containers, all proteins, functionally active fragments or functionally active derivatives thereof of from the group of proteins in column C of table 1.

The kits of the present invention can also contain expression vectors encoding the essential components of the complex machinery, which components after being expressed can be reconstituted in order to form a biologically active complex. Such a kit preferably also contains the required buffers and reagents. Optionally associated with such container(s) can be instructions for use of the kit and/or a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

5.8 ANIMAL MODELS

The present invention also provides animal models. In one embodiment, animal models for diseases and disorders involving the protein complexes of the present invention are provided. These animal models are well known in the art. These animal models include, but are not limited to those which are listed in the section 5.6 (supra) as exemplary animal models to study any of the complexes provided in the invention. Such animals can be initially produced by promoting homologous recombination or insertional mutagenesis between genes encoding the protein components of the complexes in the chromosome, and exogenous genes encoding the protein components of the complexes that have been rendered biologically inactive or deleted (preferably by insertion of a heterologous sequence, *e.g.*, an antibiotic resistance gene). In a preferred aspect, homologous recombination is carried out by transforming embryo-derived stem (ES) cells with one or more vectors containing one or more insertionally inactivated genes, such that homologous recombination occurs, followed by injecting the transformed ES cells

into a blastocyst, and implanting the blastocyst into a foster mother, followed by the birth of the chimeric animal ("knockout animal") in which a gene encoding a component protein from column A of table 1 of a given row, or a functionally active fragment or functionally active derivative thereof, and a gene encoding a component protein from column B of table 1 of said row, or a functionally active fragment or functionally active derivative thereof, has been inactivated or deleted (Capecchi, 1989, Science 244:1288-1292)..

In another preferred aspect, homologous recombination is carried out by transforming embryo-derived stem (ES) cells with one or more vectors containing one or more insertionally inactivated genes, such that homologous recombination occurs, followed by injecting the transformed ES cells into a blastocyst, and implanting the blastocyst into a foster mother, followed by the birth of the chimeric animal ("knockout animal") in which the genes of all component proteins from the group of proteins listed in column C of table 1 or of all proteins from the group of proteins listed in column D of table 1 have been inactivated or deleted.

The chimeric animal can be bred to produce additional knockout animals. Such animals can be mice, hamsters, sheep, pigs, cattle, etc., and are preferably non-human mammals. In a specific embodiment, a knockout mouse is produced.

Such knockout animals are expected to develop, or be predisposed to developing, diseases or disorders associated with mutations involving the protein complexes of the present invention, and thus, can have use as animal models of such diseases and disorders, e.g., to screen for or test molecules (e.g., potential Therapeutics) for such diseases and disorders.

In a different embodiment of the invention, transgenic animals that have incorporated and express (or over-express or mis-express) a functional gene encoding a protein component of the complex, e.g. by introducing the a gene encoding one or more of the components of the complex under the control of a heterologous promoter (*i.e.*, a promoter that is not the native promoter of the gene) that either over-expresses the protein or proteins, or expresses them in tissues not normally expressing the complexes or proteins, can have use as animal models of diseases and disorders characterized by elevated levels of the protein complexes. Such animals can be used to screen or test molecules for the ability to treat or prevent the diseases and disorders cited *supra*.

In one embodiment, the present invention provides a recombinant non-human animal in which an endogenous gene encoding a first protein, or a functionally active fragment or functionally active derivative thereof, which first protein is selected from the group of proteins of column A of table 2 of a given complex, and an endogenous gene encoding a second protein, or a functionally active fragment or functionally active derivative thereof, which second protein is selected from the group consisting of proteins of column B, of table 2 of said complex has been deleted or inactivated by homologous recombination or insertional mutagenesis of said animal or an ancestor thereof. In addition, the present invention provides a recombinant non-human animal in which the endogenous genes of all proteins, or functionally active fragments or functionally active derivatives thereof of one of the group of proteins listed in column C have been deleted or inactivated by homologous recombination or insertional mutagenesis of said animal or an ancestor thereof:

In another embodiment, the present invention provides a recombinant non-human animal in which an endogenous gene encoding a first protein, or a functionally active fragment or functionally active derivative thereof, which first protein is selected from the group consisting of proteins of column A of table 2 of a given complex, and an endogenous gene encoding a second protein, or a functionally active fragment or functionally active derivative thereof, which second protein is selected from the group consisting of proteins of column B, of table 2 of said complex are recombinantly expressed in said animal or an ancestor thereof.

The following series of examples are presented by way of illustration and not by way of limitation on the scope of the invention.

EXAMPLES

By applying the process according to the invention to the isolation of the polyadenylation/cleavage machinery from yeast, which is further described below, thirty-two new proteins could be identified in said yeast complex.

Purifications have been done using different proteins as bait according to the protocols stated further below.

Below is a more detailed list of the newly identified components of the polyadenylation complex (see also Tab. 1). The Accession-Number stated is the GenBank-Accession number for the protein.

Protein patterns for some of the purifications are shown in Figures 3 and 4.

Act1: Is a known and essential protein (GenBank Acc. No. BAA21512.1), which has been shown to be involved in Pol II transcription and has been found to be associated with histone acetylation. It serves as a structural protein.

Cka1: Is a known and non-essential protein (GenBank Acc. No. CAA86916.1), which has been found to be involved in Polymerase III transcription and has been found to be associated with the Casein kinase II complex.

Eft2: The translation elongation factor EF-2 is a known protein involved in protein synthesis (GenBank AAB64827.1)

Eno2: Is a known and essential protein (GenBank Acc. No. AAB68019.1). It has been shown to have lyase activity and is known to be involved in carbohydrate metabolism.

Glc7 (YER133w) is also a known protein (GenBank Acc. No. AAC03231.1). It is also an essential protein and is a Type I protein serine threonine phosphatase which has been implicated in distinct cellular roles, such as carbohydrate metabolism, meiosis, mitosis and cell polarity. Its occurrence in the cleavage/polyadenylation machinery has not been known before.

Gpm1: This protein is a phosphoglycerate mutase that converts 2-phosphoglycerate to 3-phosphoglycerate in glycolysis. It is an essential protein (GenBank: CAA81994.1)

Hhf2: Is a known and non-essential protein (GenBank Acc. No. CAA95892.1) which has been shown to be involved in DNA-binding. It has previously been linked to Histone octamer and the RNA polymerase I upstream activation factor.

Hta1: Is a known and non-essential protein (GenBank Acc. No. CAA88505.1) which has DNA-binding capability and has been shown to be involved in polymerase II transcription.

Hsc82: Is a non-essential protein so far being associated with protein folding. (GenBank Acc. No: CAA89919.1)

Imd2: Is an Inosine-5'-monophosphate dehydrogenase so far being associated with nucleotide metabolism. It is non-essential. (GenBank Acc.-No.: AAB69728.1)

Imd4: Is a non-essential protein with similiarity to Imd2 so far being associated with nucleotide metabolism (GenBank Acc-No.: CAA86719.1)

Met6: Is a homocysteine methyltransferase so far being associated with amino-acid metabolism (GenBank Acc.-No.: AAB64646.1)

Pdc1: Is a pyruvate decarboxylase isozyme1 so far being associated with carbohydrate metabolism (GenBank Acc.-No.: CAA97573.1)

Pfk1: Is a known protein (GenBank Acc. No. CAA97268.1) which has previously been described as part of the phosphofructokinase complex.

Ref2 (YDR195w) is a known protein (GenBank Acc. No. CAA88708.1). It is a non-essential gene product. It has been shown to be involved in 3'-end formation prior to the final polyadenylation step. However, Ref2 has never been identified before as a component of the 3'-end processing machinery. Ref2 has been shown to interact with Glc7, another new component of the cleavage/polyadenylation machinery.

Sec13: Is a known and essential protein (GenBank Acc. No AAB67426.1).

Sec31: Is a known and essential protein (GenBank Acc. No. CAA98772.1)

Ssa3: Is a known and non-essential protein (GenBank Acc. No. CAA84896.1) which so far has been implicated with protein folding/protein transport.

Ssu72 (YNL222w) is also a known protein (GenBank Acc. No. CAA96125.1) and is an essential phylogenetically conserved protein which has been shown to interact with the general transcription factor TFIIB (Sua7). TFIIB is an essential component of the RNA polymerase II (RNAP II) core transcriptional machinery. It is thought that this interaction plays a role in the mechanism of start site selection by RNAP II. The finding according to the present invention that Ssu72 is associated with Pta1 is likely to be relevant since it is believed that mRNA 3'-end formation is linked with other nuclear processes like transcription, capping and splicing. Furthermore, Ssu 72 has also been clearly identified in a "reverse tagging experiment" as explained herein below by using some of the Pta1 associated proteins as bait. However, when Ssu72 itself was used as a bait associated proteins were not found most likely due to the fact that the addition of a C-terminal tag renders Ssu72 non-functional.

Taf60: Is a known and essential protein (GenBank Acc. No. CAA96819.1) which has been shown to be involved in Polymerase II transcription.

Tkl1: Is a non-essential transketolase so far being associated with amino-acid metabolism and carbohydrate metabolism (GenBank Acc-No.: CAA89191.1)

Tsa1: Translation initiation factor eIF5 which so far has been to shown to catalyze hydrolysis of GTP on the 40S ribosomal subunit-initiation complex followed by joining to 60S ribosomal subunit. (GenBank Acc.-No.: CAA92145.1)

Tye7: Is a known protein (GenBank Acc. No. CAA99671.1). It has been shown to be a basic helix-loop-helix transcription factor.

Vid24: Is a known and non-essential protein (GenBank Acc. No. CAA89320.1) which has previously been associated with protein degradation and vesicular transport.

Vps53: Is a known protein (GenBank Acc. No. CAA89320.1) which has been found to play a role in protein sorting.

YCL046w: Is a non-essential protein (GenBank Acc. No. CAA42371.1).

YGR156w is the protein product of an essential gene. This protein also contains a RNA binding motif. (GenBank Acc. No. CAA97170.1).

YHL035c: Is a known and non-essential protein (GenBank Acc. No. AAB65047.1). It is a member of the ATP-binding cassette superfamily.

YKL018w is also an essential protein containing a WD40 domain which is a typical protein binding domain. (GenBank Acc. No. CAA81853.1)

YLR221c: Is a protein of unknown function (GenBank Acc. No. AAB67410.1)

YML030w: Is a protein of unknown function (GenBank Acc. No. CAA86625.1)

YOR179c shows significant sequence similarity to Ysh1 (GenBank Acc. No. CAA99388.1)

Two further proteins for which binary interactions with members of the polyadenylation complex as known so far have been shown before have also been purified with the complex:

YKL059c: is the product of an essential gene and is a zinc binding protein containing a C2HC Zinc finger. The presence of this domain predicts a RNA binding function of YKL059c. We believe the corresponding gene product is identical to Pfs1, a protein which has been mentioned in several publications, but which has never been annotated in the databases (for review see Keller, W. and Minvielle-Sebastia (1997). Curr Opin Cell Biol 11: 352-357). (GenBank Acc. No. CAA81896.1)

Tif4632: Is a known and non-essential protein (GenBank Acc. No. CAA96751.1) which has been shown to have an RNA-binding/translation factor activity and is involved in protein synthesis.

Below is a description of the experimental steps and protocols as used herein:

The initial round of purification of the complex was carried out using Pta-1 as a bait as described below:

CONSTRUCTION OF A YEAST STRAIN EXPRESSING TAP-TAGGED Pta1

The construction of these strains is illustrated both in Figure 2 and table 5.

PURIFICATION OF PROTEINS ASSOCIATED WITH PTA1

The TAP-technology, which is more fully described in WO/0009716 and in Rigaut, G. et. al. (1999), Nature Biotechnology. Vol. 17 (10): 1030-1032 respectively was used for protein complex purification. The Pta1 protein was C-terminally tagged with a TAP-tag which consists of calmodulin-binding peptide (CBP), a cleavage site for TEV protease followed by two IgG-binding units of protein A (Rigaut, G. et. al. (1999), Nature Biotechnology. Vol. 17 (10): 1030-1032). Pta1 is an essential protein which has been reported to be a component of PFI. Pta1-TAP was used as a bait to identify associated partners from cell lysates using the two-step TAP purification procedure. Proteins were separated by 1D gel electrophoresis and visualized by staining with Coomassie. More than a total of 20 bands could be detected on the gel (see Fig. 3). The identity of the proteins was determined by mass spectrometry. 13 of these are known components of the pre-mRNA processing machinery: Cft1, Cft2, Ysh, Pta1, Rna14, Pab1, Pcf11, Pap1, Clp1, Pfs2, Fip1, Rna15 and Yth1. It is to be noted that such a comprehensive number has never before been purified together in form of a complex. The remaining seven proteins have not previously been found associated with Pta1: Ref2, YK059c, YGR156w, YKL018w, Glc7, Ssu72 and YOR179c.

VALIDATION OF INTERACTIONS FOUND WITH Pta1

A reciprocal experiment to the one described above was performed. For this purpose a subset of the interactors found in the above described Pta purification (both known and novel interactors) were chosen as a bait for a further round of purification (the baits used herein are listed in the first column of Table 1). In the case of some proteins the C-terminally tagged versions could not be recovered. The likely reason for this is that the addition of the TAP tag at the C-terminus interferes with the function of these proteins.

An important fact is that almost all of the known components involved in 3'-end formation and five of the seven novel proteins identified herein are essential for cell viability. The protein pattern obtained in some of those experiments is shown in Figure 4.

The construction of the strains was carried out as described for the strain expressing the TAP-tagged Pta-1.

SEQUENCE ANALYSIS OF MEMBERS OF THE COMPLEX

The process of mRNA processing is highly conserved in eukaryotes. Accordingly, for a number of the yeast proteins human orthologues could be found (see Table 2). This illustrates that many of the functions found in the yeast complex can be transferred to humans. Also the enzymatic activity of this complex has long been known, the enzymatically active member could not yet be unraveled. Using extensive sequence similarity searches it could be shown that Ysh1 is homologous to a class of bacterial beta-lactamases. The active center of this protein family contains 2 zinc ions which are bound by histidines. As these residues are conserved in Ysh1 and it was shown that enzymatic activity of the yeast complex is zinc dependent predicted that Ysh1 is responsible for the catalytic activity of the complex. Two other proteins found in the complex, Cft2 and YOR179c, are homologous to the Ysh1 N- and C-terminus, respectively. Though Cft2 is homologous to the enzymatic region of Ysh1 it misses the zinc binding histidines indicating that it lacks enzymatic activity. Thus, Cft 2 and YOR179c could compete with Ysh1 for the same binding slot of the complex, suggesting a novel type of regulation of polyadenylation. A similar way of regulation might be used in the case of Pfs2 and YKL018w, which both consist of multiple WD40 domains.

PREDICTION OF MAMMALIAN PROTEINS

To allow the transfer of function information from yeast to human proteins, we did not only use an identity cutoff, but also the 'orthology' concept. Orthology defines genes which arose via a speciation event, in contrast to genes which arose via gene duplication. Orthologue genes are supposed to perform the same function in different organisms, therefore more detailed function information can be transferred. The algorithm for the detection of orthologous gene pairs from yeast and human uses the

whole genome of these organisms. First, pairwise best hits were retrieved, using a full Smith-Waterman alignment of predicted proteins. To further improve reliability, these pairs were clustered with pairwise best hits involving *Drosophila melanogaster* and *Caenorhabditis elegans* proteins. See "Initial sequencing and analysis of the human genome", Nature 2001 Feb 15; 409(6822):860-921 for a detailed description of the analysis.

Bioinformatic analysis of the Complex:

Functional domains of all members of the complex were analyzed using SMART (SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res 2000 Jan 1; 28(1):231-4) and Pfam (Pfam: protein families database, Nucleic Acids Res 2000 Jan 1; 28(1):263-6).

COMPARISON OF THE YEAST AND MAMMALIAN CLEAVAGE/POLYADENYLATION MACHINERY

The sequence of many of the polypeptides involved in 3'-end formation are conserved from yeast to mammals, although the sequence elements on the substrate pre-mRNA differ (see Figure 1).

The detailed experimental protocols for the example stated herein are given below:

PROTOCOLS:

ISOLATION OF PROTEIN COMPLEXES:

a) ISOLATION OF COMPLEXES FROM YEAST:

Yeast strain construction:

Yeast strains expressing TAP-tagged ORFs were constructed in a semi-automated way essentially according to Rigaut et. al. (Rigaut, G. et. al. Nat Biotechnol 17, 1030-2 (1999)) and Puig et al. (Puig, O. et al. Methods 24, 218-19. (2001)) (See also Fig. 2 and Table 5)

TAP-purification using the Pta-1-tagged strains::

Pta1-tagged strain was cultured in 4 l of YPD medium to an OD600 of 2.

After harvesting, the cell pellet was frozen in liquid nitrogen and stored at -80°C. All further manipulations were done at 4°C except where noted. For preparation of protein lysates the cells were resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.15 % NP-40, 1.5 mM MgCl₂, 0.5 mM DTT, protease inhibitors) and subjected to mechanical disruption with glass beads. Lysates were clarified by two successive centrifugation steps at 20.000 x g for 10 min and 100.000 x g for 1 hour. After addition of glycerol to 5 % final concentration the lysates were frozen in liquid nitrogen and stored at -80°C.

For the first purification step 500 µl of rabbit IgG-Agarose (50:50 slurry, Sigma A2909) pre-equilibrated in lysis buffer were added to the lysate and the sample was rotated for 2 hours. The unbound fraction was discarded and the beads with the bound material were transferred to a 0.8 ml column (MoBiTec M1002, 90 µm filter). The beads were washed with 10 ml of lysis buffer followed by 5 ml of TEV cleavage buffer (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.1 % NP-40, 0.5 mM EDTA, 1 mM DTT).

150 µl of TEV cleavage buffer and 4 µl of TEV protease were added to the column and the sample was incubated on a shaker at 16 °C for 2 hours. The eluate was recovered by pressing with a syringe.

150 µl of Calmodulin dilution buffer (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.1 % NP-40, 2 mM MgAc, 2 mM imidazole, 4 mM CaCl₂, 1 mM DTT) was added to the previous eluate and this mixture was transferred to a MoBiTec column containing 300 µl (bead volume) of Calmodulin affinity resin (Stratagene #214303) which was prewashed in Calmodulin wash buffer (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.1 % NP-40, 1 mM MgAc, 1 mM imidazole, 2 mM CaCl₂, 1 mM DTT). The samples were rotated for 1 hour at 4 °C.

After washing of the beads with 10 ml of Calmodulin wash buffer, protein complexes were eluted with 600 μ l of elution buffer (10 mM Tris-HCl pH 8.0, 5 mM EDTA). The samples were concentrated in siliconised tubes in a speed vac to a final volume of 10-20 μ l. Proteins were detected by polyacrylamide gel electrophoresis followed by staining with colloidal Coomassie blue.

General TAP-purification protocol for soluble proteins:

TAP-purification of soluble proteins:

The purification was done from 2 liters of yeast cells grown to late log phase (OD_{600} ~3 – 4). Cells were harvested and the pellet was frozen in liquid nitrogen and stored at -80°C . All steps were done at 4°C . For preparation of protein lysates the cells were resuspended in buffer A (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.15 % NP-40, 1.5 mM MgCl_2 , 0.5 mM DTT, protease inhibitors) and subjected to mechanical disruption with glass beads. Lysates were clarified by two successive centrifugation steps at $20,000 \times g$ for 10 min and $100,000 \times g$ for 1 hour. After addition of glycerol to 5 % final concentration the lysates were frozen in liquid nitrogen and stored at -80°C .

For the first purification step 500 μ l of rabbit IgG-Agarose (50:50 slurry, Sigma A2909) pre-equilibrated in buffer A were added to the lysate and the sample was rotated for 1 hour. The unbound fraction was discarded and the beads with the bound material were transferred to a 0.8 ml column (MoBiTec M1002, 90 μ m filter). The beads were washed with 10 ml of buffer A.

150 μ l of buffer A and 4 μ l of TEV protease (1 mg/ml) were added to the column and the sample was incubated on a shaker at 16°C for 1 hour. The eluate was recovered by pressing with a syringe.

150 μ l of buffer A containing 4 mM CaCl_2 was added to the previous eluate and this mixture was transferred to a MoBiTec column containing 300 μ l (bead volume) of Calmodulin affinity resin (Stratagene #214303) which was prewashed in buffer A containing 2 mM CaCl_2 . The samples were rotated for 1 hour at 4°C .

After washing of the beads with 5 ml of buffer A containing 2 mM CaCl_2 , protein complexes were eluted with 600 μ l of elution buffer (10 mM Tris-HCl pH 8.0, 5 mM EGTA). The samples were concentrated in siliconized tubes in a speed vac. Proteins

were detected by polyacrylamide gel electrophoresis followed by staining with colloidal Coomassie blue.

TAP-purification of membrane proteins:

The purification was done from 2 liters of yeast cells grown to late log phase ($OD_{600} \sim 3 - 4$). Cells were harvested and the pellet was frozen in liquid nitrogen and stored at -80°C . All steps were done at 4°C . For the purification of TAP-tagged membrane proteins cells were lysed in buffer containing 50 mM Hepes/KOH pH 7.5, 150 mM KCl, 0.25 % NP-40, 2 mM MgCl_2 , 2 mM EDTA, 0.5 mM DTT and protease inhibitors. The extracts were spun at $20,000 \times g$ for 10 min and the resulting supernatant was adjusted to 1.5 % NP-40 and 5 % glycerol. Samples were incubated for 30 min with end-over-end shaking and then centrifuged at $180,000 \times g$ for 30 min. The resulting supernatant was immediately used for TAP-purification.

For the first purification step 500 μl of rabbit IgG-Agarose (50:50 slurry, Sigma A2909) pre-equilibrated in buffer B (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.5 % NP-40, 1.5 mM MgCl_2 , 0.5 mM DTT, protease inhibitors) was added to the lysate and the sample was rotated for 1 hour. The unbound fraction was discarded and the beads with the bound material were transferred to a 0.8 ml column (MoBiTec M1002, 90 μm filter). The beads were washed with 10 ml of buffer B.

150 μl of buffer B and 8 μl of TEV protease (1 mg/ml) were added to the column and the sample was incubated on a shaker at 16°C for 1 hour. The eluate was recovered by pressing with a syringe.

150 μl of buffer B containing 4 mM CaCl_2 was added to the previous eluate and this mixture was transferred to a MoBiTec column containing 300 μl (bead volume) of Calmodulin affinity resin (Stratagene #214303) which was prewashed in buffer B containing 2 mM CaCl_2 . The samples were rotated for 1 hour at 4°C .

After washing of the beads with 5 ml of buffer B containing 2 mM CaCl_2 , protein complexes were eluted with 600 μl of elution buffer (10 mM Tris-HCl pH 8.0, 5 mM EGTA). The samples were concentrated in siliconized tubes in a speed vac. Proteins were detected by polyacrylamide gel electrophoresis followed by staining with colloidal Coomassie blue.

b) ISOLATION OF COMPLEXES FROM MAMMALIAN CELLS

ISOLATION OF COMPLEXES FROM MAMMALIAN CELLS

Cells:

Retroviral transduction vectors containing the TAP-cassette were generated by directional cloning of PCR-amplified ORFs into a modified version of a MmoLV-based vector via the Gateway site-specific recombination system (Life Technologies). Virus stocks were generated in a HEK293 gag-pol packaging cell line by pseudotyping with VSV-G. Cells were infected and complexes were purified after cell expansion and cultivation of 3-5 using a modified TAP-protocol

Standard lysis protocol:

The medium was removed from the culture dish and the cells were scraped directly from the plate with help of a rubber policeman. The cells were collected on ice washed 3 times with PBS and resuspended in lysis buffer (50 mM Tris, pH: 7.5; 5 % glycerol; 0,2 % IGEPAL; 1.5 mM MgCl₂; 1 mM DTT; 100 mM NaCl; 50 mM NaF; 1 mM Na₃VO₄ + protease inhibitors). The cells were lysed for 30 min on ice, spun for 10 min. at 20,000g and re-spun for 1h at 100,000g. The supernatant was recovered, rapidly frozen in liquid nitrogen and stored at -80 °C. For pre-clearing the thawed lysate was incubated with 500 µl sepharose CL-4B beads (Amersham Pharmacia) for 1 h shaking and finally processed according the TAP protocol.

Nuclear lysis protocol:

The medium was removed from the culture dish and the cells were scraped directly from the plate with help of a rubber policeman. The cells were collected on ice washed 3 times with PBS and resuspended in buffer A (10 mM Tris-Cl, pH 7.5; 1, 5 mM MgCl₂; 10 mM KCl; 1 mM DTT, 50 mM NaF; 1 mM Na₃VO₄). To isolate the nuclei the lysate was dounced with a tight fitted pestle in a dounce homogenizer for 15 strokes. The nuclei were harvested by centrifugation (10 min. at 2000 g and 20 min. at 16,000 g) and lysed in

buffer B (50 mM Tris-Cl, pH: 7.5; 1.5 mM MgCl; 20 % glycerol; 420 mM NaCl; 1mM DTT; 50 mM NaF; 1 mM Na₃VO₄) for 30 min. on ice with frequent shaking. The protein lysate was cleared by centrifugation (30 min at 100,000 g) and 1 : 4 diluted with buffer C (50 mM Tris-Cl, pH: 7.5; 1 mM DTT; 0.26 % NP40; 1.5 mM MgCl; 50 mM NaF; 1 mM Na₃VO₄). After 30 min incubation on ice the lysate was re-spun for 30 min at 100,000 g, quickly frozen in liquid nitrogen and stored at -80 °C. For pre-clearing the thawed lysate was incubated with 500 µl sepharose CL-4B beads (Amersham Pharmacia) for 1 h shaking and finally processed according the TAP protocol.

MASS SPECTROMETRIC ANALYSIS

Protein digestion prior to mass spectrometric analysis:

Gel-separated proteins were reduced, alkylated and digested in gel essentially following the procedure described by Shevchenko et al. (Shevchenko, A., Wilm, M., Vorm, O., Mann, M. Anal Chem 1996, 68, 850-858). Briefly, gel-separated proteins were excised from the gel using a clean scalpel, reduced using 10 mM DTT (in 5mM ammonium bicarbonate, 54 °C, 45 min) and subsequently alkylated with 55 mM iodoacetamid (in 5 mM ammonium bicarbonate) at room temperature in the dark (30 min). Reduced and alkylated proteins were digested in gel with porcine trypsin (Promega) at a protease concentration of 12.5 ng/µl in 5mM ammonium bicarbonate. Digestion was allowed to proceed for 4 hours at 37 °C and the reaction was subsequently stopped using 2 µl 25% TFA.

Desalting and concentration of peptides produced by in-gel digestion of gel-separated proteins:

Peptides were desalted and concentrated using a prefabricated uZipTip (Millipore) reversed phase column. Peptides were eluted directly onto stainless steel MS sample holders using 2µl eluent (70% acetonitrile in 5% TFA containing 2mg/ml alpha-Cyano-4-hydroxy-cinnamic acid and two standard peptides for internal calibration of mass spectra).

Mass spectrometric data acquisition:

Matrix-assisted laser desorption/ionisation (MALDI) time-of-flight (TOF) mass spectra were acquired in delayed extraction mode on a Voyager DE-STR PRO MALDI mass spectrometer (Applied Biosystems) equipped with a 337 nm nitrogen laser. 500 laser shots were averaged in order to produce final spectra. Spectra were automatically internally calibrated using two standard peptides. The monoisotopic masses for all peptide ion signals detected in the acquired spectra were determined and used for database searching.

Protein sequence database searching using peptide mass fingerprinting (PMF) data: The list of monoisotopic peptide masses obtained from the MALDI mass spectrum was used to query a fasta formatted protein sequence database that contained all protein sequences from *S. cerevisiae*. Proteins were identified by peptide mass fingerprinting (Mann, M., Højrup, P., Roepstorff, P. *Biol Mass Spectrom* 1993, 22, 338-345; Pappin, D., Højrup, P., Bleasby, A. J. *Curr. Biol.* 1993, 3, 327-33; Henzel, W. J., Billeci, T. M., Stults, J. T., Wong, S. C., Grimley, C., Watanabe, C. *Proc Natl Acad Sci U S A* 1993, 90, 5011-5015; Yates, J. R., Speicher, S., Griffin, P. R., Hunkapiller, T. *Anal Biochem* 1993, 214, 397-408; James, P., Quadroni, M., Carafoli, E., Gonnet, G. *Biochem Biophys Res Commun* 1993, 195, 58-64) using the software tool ProFound (Proteometrics). In PMF, a protein is identified by correlating the measured peptide masses with theoretical digests of all proteins present in the database. Search criteria included: tryptic protein cleavage, monoisotopic masses, 30 ppm mass accuracy. No restrictions on protein size or isoelectric point were imposed.

BIOINFORMATICS

Functional and localization information about yeast proteins was retrieved from the Yeast Protein Database (YPD (Constanzo, M.C. et al., 2001, *Nucl. Acid Res*, 29: 75-9; Hodges, P.E. et al., 1999, *Nucl. Acids Res* 27: 69-73)) released in August 2001. In order to get a more concise classification for localization and function, YPD classes were merged. Protein domain analysis was performed using SMART (Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P. & Bork, P. Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P. & Bork, P. SMART, *Nucleic Acids Res* 28, 231-4. (2000)). PsiBlast (Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: *Nucleic Acids Res* 25, 3389-402. (1997)) was used for

homology analysis. All additional analysis software has been developed in house, using Perl and Python.

ASSAYS FOR ASSAYING THE ACTIVITIES OF THE COMPLEXES PRESENTED IN THE INVENTION

An exemplary RNA binding assay can be carried out by contacting a complex having RNA binding activity with a radioactive [³²P] end-labeled RNA substrate, e.g. a poly (A) RNA, under appropriate conditions and detecting bound protein. The detection of bound protein can be carried out, e.g., by filtrating the solution through a nitrocellulose filter and determining the radioactivity bound to the filter. This assay is based on the retention of nucleic acid-protein complexes on Nitrocellulose whereas free nucleic acid can pass through the filter

(see e.g. Wahle, E., 1991, Methods 66: 759-68)

An exemplary RNA exonuclease assay can be carried out by contacting a complex having RNA exonuclease activity with a radioactivity [³² phosphate] end-labeled RNA substrate under appropriate conditions and detecting the release of free radioactive nucleotides. The detection of free radioactive nucleotides can be carried out, e.g., by adding 20% trichloroacetic acid, filtrating the solution through a filter and measuring the amount of acid-soluble radioactivity

(see e.g. Ross, J., 1999, Methods 17: 52-9)

An exemplary mRNA splicing assay can be carried out by contacting a complex having mRNA splicing activity with a radioactively labeled RNA substrate under appropriate conditions and detecting the release of spliced RNA species. The detection of spliced RNA species can be carried out, e.g., by fractionation of processed RNAs in a glycerol gradient and subsequent analysis by denaturing polyacrylamide gel electrophoresis and visualization by autoradiography.

(see e.g. Schwer, B. and Gross, CH., 1998, Methods 17: 2086-94)

An exemplary rRNA processing assay can be carried out by contacting a complex having rRNA processing activity with an pre-rRNA substrate under appropriate conditions and

detecting the release of free processed rRNA species. The detection of processed rRNA species can be carried out, e.g., using a primer extension or northern blotting assay by measuring the size of the rRNA species.

(see e.g. Kressler, D. et al, 1997, Methods 17: 7283-94)

TABLE 1

Entry point	Interactions	Proteins found	COLUMN A: known components of the complex	COLUMN B: novel proteins	COLUMN C: Cleavage/poly-adenylation machinery	Activity of the complex	Column D: Proteins of unknown function	local.
Cft1	Cft1, Cft2, Pta1, Ysh1	Act1 Cft1	Cft1 Cft2	Act1 Cka1	Act1 Cft1 Cft2	3' end	Ycl046w	c n u
Cft2	coippt; Cft1, Cft2,	Cft2 Cka1	Clp1 Fip1	Eft2 Eno2	Cka1 Clp1 Eft2	processing	Ygr156w	e
Clp1	Fip1, Pap1, Pta1,	Clp1 Eft2	Pab1 Pap1	Glc7 Gpm1	Eno2 Fip1 Glc7	activity for	Yhl035c	
Fip1	Ykl059c, Ysh1, Yth1	Eno2 Fip1	Pcf11 Pfs2	Hhf2 Hta1	Gpm1 Hhf2	mRNA	Ykl018w	
Pap1	coippt; Fip1, Yth1 2-	Glc7	Pta1 Rna14	Hsc82 Imd2	Hsc82 Hta1		Ylr221c	
Pcf11	hybrid, coippt, high-	Gpm1	Rna15	Imd4 Met6	Imd2 Imd4		Yml030w	
Pfs2	throughput 2-hybrid, in	Hhf2	Tif4632	Pdc1 Pfk1	Met6 Pab1		Yor179c	
Pta1	vitro binding Glc7,	Hsc82	Ykl059c Ysh1	Ref2 Sec13	Pap1 Pcf11			
Ref2	Ref2; Fip1, Rna14 2-	Hta1	Yth1	Sec31	Pdc1 Pfk1 Pfs2			
Rna14	hybrid; Rna14, Rna15	Imd2		Ssa3	Pta1 Ref2			
Rna15	2-hybrid; Clp1, Pcf11	Imd4		Ssu72	Rna14 Rna15			
Ykl059c	coippt, high-	Met6		Taf60 Tkl1	Sec13 Sec31			
Yor179c	throughput, 2-hybrid;	Pab1		Tsa1 Tye7	Ssa3 Ssu72			
Ysh1	Fip1, Pap1 2-hybrid,	Pap1		Vid24	Taf60 Tif4632			
Yth1	coippt; Fip1, Ysh1	Pcf11		Vps53	Tkl1 Tsa1 TYe7			
	coippt, copurification;	Pdc1 Pfk1		Ycl046w	Vid24 Vps53			
	Hhf2, Hta1 affinity	Pfs2 Pta1		Ygr156w	Ycl046w			
	column, coippt;	Ref2		Yhl035c	Ygr156w			

Sec13, Sec31 2-hybrid, affinity column, Rna14	Rna14			Yhl035c		
copurification, high-throughput 2-hybrid; Pab1, Pcf11p, Rna14p, Rna15p	Rna15		Ykl018w	Ykl018w		
copurification, Pab1, Tif4632	Sec13		Ylr221c	Ykl059c		
Tif4632 copurification, in vitro binding; Pcf11, Tkl1 Tsa1	Sec31		Yml030w	Ylr221c		
Rna14, Rna15	Ssa3		Yor179c	Yml030w		
copurification; Pcf11, Rna14 2-hybrid,	Ssu72			Yor179c Ysh1		
coippt, copurification, high-throughput 2-hybrid; Pcf11, Rna15	Taf60			Yth1		
2-hybrid,	Tif4632					
copurification, high-throughput 2-hybrid;	Tkl1 Tsa1					
Pfs2, Rna14 2-hybrid, coippt; Pfs2, Ysh1	TYe7					
coippt; Fip1, Pfs2	Vid24					
coippt; Pap1, Yth1	Vps53					
coippt; Clp1, Rna14	Ycl046w					
	Ygr156w					
	Yhl035c					
	Ykl018w					
	Ykl059c					
	Ylr221c					
	Yml030w					
	Yor179c					
	Ysh1					
	Yth1					

TABLE 2

INDIVIDUAL YEAST PROTEINS OF THE COMPLEXES

A)

Yeast proteins listed in table 2	SEQ ID	MIPS	SWISS-PROT	SGD	Genbank	human orthologue in GenBank	C.elegans orthologue in GenBank	Drosophila orthologue in GenBank
ACT1	1	ATBY	P02579	S0001855	D50617	gi4501885	ACT4_CAEEL L	Q9W460
CFT1	3	S61187		S0002709	U28374			
CFT2	5	S64952		S0004105	Z73287			
CKA1	7	A31564	P15790	S0001297	Z46861			
CLP1	9	S67147		S0005776	Z75158	gi5803029	YMI4_CAEEL L	Q9V6Q1
EFT2	11		P32324	S0002793	AAB64827.1	gi:4503483	EF2_CAEEL	Q9V9R0
ENO2	13	NOBY2	P00925	S0001217	U00027	gi:4503571	ENO_CAEEL L	Q9VQ38

GLC7	15	S32595	P32598	S0000935	U18916	gi4506003	YME1_CAE EL	Q9VC69
GPM1	17	PMBYY	P00950	S0001635	Z28152			
Fip1	19	A56545	P45976	S0003853	Z49593			
HHF2	21		P02309	S0004975	Z71306			
HTA1	23	HSBYA1	P04911	S0002633	Z48612			
HSC82	25	S55133	P15108	S0004798	CAA89919.1	gi6680307	gi3875041	gi7292327
IMD2	27	S48997	P38697	S0001259	AAB69728.1	gi4504689	gi18030187	gi7291188
IMD4	29	S50890	P50094	S0004520	CAA86719.1	gi4504688	gi18030187	gi7291188
MET6	31	S50594	P05694	S0000893	AAB64646.1			
PAB1	33	DNBYPA	P04147	S0000967	U18922	F15P000000 84557	Q9U302	Q9V8C3
PAP1	35	S19031	P29468	S0001710	Z28227	P51003	Q20370	Q9V8X7
PCF11	37	S59435	P39081	S0002636	Z48612	gi7706224	YRR2_CAE EL	Q9V768
PDC1	39	DCBYP	P06169	S0004034	Z73216			
PFK1	41	JQ0016	P16861	S0003472	Z73025			

YSH1	75	S51413		S0004267	U17245			
YTH1	77	S59772		S0006311	U32445			
YCL046W	79	S19375	P25575	S0000551	X59720			
YGR156	81	S60446	P39927	S0003388				
W								
YHL035C	83	S48933	P38735	S0001027	U11583			
YKL018W	85	S37831	P36104	S0001501	Z28018	F15P000000	Q18403	Q9VLN1
						64108		
YLR221C	87	S51444		S0004211	U19027			
YKL059C	89	S37881	P35728	S0001542	Z28059			
YML030	91	S49749	Q03713	S0004492	Z46659			
W								
YOR179C	93	S67071		S0005705	Z75087			

B)

Yeast proteins listed in table 2	YPD description	Biochemical function from YPD	Cellular function from YPD
ACT1	Actin, involved in cell polarization, endocytosis, and other cytoskeletal functions	Structural protein	Cell polarity; Cell structure; Chromatin/chromosome structure; Mating response; Pol II transcription; Vesicular transport
CFT1	Component of pre-mRNA cleavage factor II	Hydrolase; Nuclease [endo, exo, ribo, deoxyribo]	RNA processing/modification
CFT2	Component of pre-mRNA cleavage factor II	RNA-binding protein	RNA processing/modification
CKA1	Casein kinase II (protein kinase CK2), catalytic (alpha) subunit	Protein kinase; Transferase	Pol III transcription
CLP1	Subunit of cleavage and polyadenylation factor IA, required for 3'-end processing of pre-mRNA	Nuclease [endo, exo, ribo, deoxyribo]	RNA processing/modification
EFT2	Translation elongation factor EF-2, identical to <u>Eft1p</u> , contains diphthamide which is not essential for its activity	Translation factor	Protein synthesis

ENO2	Enolase 2 (2-phosphoglycerate dehydratase); converts 2-phospho-D-glycerate to phosphoenolpyruvate in glycolysis	Lyase	Carbohydrate metabolism
FIP1	Component of polyadenylation factor that interacts with poly(A) polymerase	RNA polymerase subunit; RNA-binding protein; Regulatory subunit	RNA processing/modification
GLC7	Protein serine/threonine phosphatase PP1 required for glucose repression, membrane bilayer mixing, and ER-to-Golgi and endocytic vesicular trafficking	Hydrolase; Protein phosphatase	Carbohydrate metabolism; Cell polarity; Cell stress; Meiosis; Mitosis
GPM1	Phosphoglycerate mutase that converts 2-phosphoglycerate to 3-phosphoglycerate in glycolysis	Isomerase	Carbohydrate metabolism; Energy generation
HHF2	Histone H4, identical to Hhf1p	DNA-binding protein	Chromatin/chromosome structure; Pol I transcription
HTA1	Histone H2A, nearly identical to Hta2p	DNA-binding protein	Cell stress; Chromatin/chromosome structure; Pol II transcription
HSC82	Chaperonin homologous to E. coli HtpG and mammalian HSP90	Heat shock protein; Hydrolase; ATPase;	Protein folding; Cell stress

		Chaperones	
IMD2	Inosine-5'-monophosphate dehydrogenase, catalyzes the conversion of inosine 5'-phosphate and NAD(+) to xanthosine 5'-phosphate and NADH, the first reaction unique to GMP biosynthesis	Oxidoreductase	Nucleotide metabolism
IMD4	Protein with similarity to inosine-5'-monophosphate dehydrogenase	Oxidoreductase	Nucleotide metabolism
MET6	Homocysteine methyltransferase (5-methyltetrahydropteroyl triglutamate-homocysteine methyltransferase), methionine synthase, cobalamin-independent	Transferase	Amino-acid metabolism
PAB1	Poly(A)-binding protein of cytoplasm and nucleus, part of the 3'-end RNA-processing complex (cleavage factor I), has 4 RNA recognition (RRM) domains	RNA-binding protein; Translation factor	Protein synthesis; RNA processing/modification; RNA turnover
PAP1	Poly(A) polymerase, required for mRNA 3' end formation, has a poorly conserved RNA recognition (RRM) domain	RNA polymerase subunit; RNA-binding protein; Transferase	RNA processing/modification
PCF11	Component of pre-mRNA cleavage and polyadenylation factor I, interacts with Rna14p and Rna15p	Nuclease [endo, exo, ribo, deoxyribo]; RNA-binding protein	RNA processing/modification
PDC1	Pyruvate decarboxylase isozyme 1	Lyase	Carbohydrate metabolism

PFK1	Phosphofructokinase alpha subunit, part of a complex with Pfk2p which catalyze ATP-dependent conversion of fructose-6-phosphate to fructose-1,6-bisphosphate, a key regulatory step in glycolysis	Other kinase; Transferase	Carbohydrate metabolism
PFS2	Polyadenylation factor I subunit 2 required for mRNA 3'-end processing, bridges two mRNA 3'-end processing factors, has WD (WD-40) repeats	Unknown	Protein complex assembly; RNA processing/modification
PTA1	Component of pre-mRNA cleavage factor II (CFII), required for both cleavage and polyadenylation of mRNA precursor	Hydrolase; Nuclease [endo, exo, ribo, deoxyribo]	RNA processing/modification
REF2	Protein involved in mRNA 3'-end formation before polyadenylation, mutant displays significantly lower usage of weak poly(A) sites	RNA-binding protein	RNA processing/modification
RNA14	Component of pre-mRNA cleavage and polyadenylation factor I (CFI) involved in poly(A) site choice, interacts with Rna15p, Fip1p, Pap1p, and Pcf11p	Hydrolase; Nuclease [endo, exo, ribo, deoxyribo]	RNA processing/modification
RNA15	Component of pre-mRNA cleavage and polyadenylation factor I (CFI), involved in poly(A) site choice, interacts with Rna14p, Pap1p, and Pcf11p, contains one RNA recognition (RRM) domain	Nuclease [endo, exo, ribo, deoxyribo]; RNA-binding protein	RNA processing/modification
SEC13	Component of the COPII coat of vesicles involved in endoplasmic reticulum to Golgi transport, contains six WD (WD-40) repeats	Unknown	Small molecule transport; Vesicular transport

SEC31	Component (p150) of the COPII coat of secretory pathway vesicles involved in endoplasmic reticulum to Golgi transport, associated with Sec13p, member of WD (WD-40) repeat family	Vesicle coat protein	Vesicular transport
SSA3	Chaperone of the HSP70 family, heat-induced cytoplasmic form not expressed under optimal conditions	Chaperones; Heat shock protein	Cell stress; Protein folding; Protein translocation
SSU72	Protein that interacts with TFIIB (Sua7p) and influences RNA polymerase II start-site selection in sua7 mutants	Complex assembly protein	Pol II transcription
TAF60	Component of TAF(II) complex (TBP-associated protein complex) and SAGA complex (Spt-Ada-Gcn5-acetyltransferase), required for activated transcription by RNA polymerase II	Transcription factor	Pol II transcription
TIF4632	mRNA cap-binding protein (eIF4F) 130K subunit	RNA-binding protein; Translation factor	Protein synthesis
TKL1	Transketolase 1	Transferase	Amino-acid metabolism; Carbohydrate metabolism
TSA1	Thioredoxin peroxidase, abundant thiol-specific antioxidant protein that prevents formation of sulfur-containing radicals	Oxidoreductase	Cell stress
TYE7	Basic helix-loop-helix transcription factor that can suppress the Gcr1p requirement for glycolytic gene expression	Transcription factor	Pol II transcription

VPS53	Subunit of the Vps52p-Vps53p-Vps54p complex, involved in protein sorting in the late Golgi	Docking protein	Unknown	Protein degradation; Vesicular transport
YCL046W	Protein of unknown function	Unknown	Unknown	Unknown
YGR156W	Protein of unknown function	Unknown	Unknown	Unknown
YHL035C	Member of the ATP-binding cassette (ABC) superfamily	ATP-binding cassette; ATPase; Active transporter, primary; Hydrolase; Transporter	ATP-binding cassette; ATPase; Active transporter, primary; Hydrolase; Transporter	Small molecule transport
YKL018W	Protein of unknown function	Unknown	Unknown	Unknown
YLR221C	Protein of unknown function	Unknown	Unknown	Unknown
YKL059C	Protein with similarity to members of the chaperonin-containing T-complex	Unknown	Unknown	Unknown
YML030W	Protein of unknown function, may be involved in mitochondrial translation	Unknown	Unknown	Energy generation
YOR179C	Protein with similarity to Ysh1p	Unknown	Unknown	Unknown
YSH1	Component of pre-mRNA cleavage factor II (CFII), required for processing of mRNA 3' end	RNA-binding protein	RNA-binding protein	RNA processing/modification
YTH1	Component of polyadenylation factor, required for both cleavage and polyadenylation of pre-mRNA	Unknown	Unknown	RNA processing/modification

TABLE 3

MEDICAL APPLICATION OF THE COMPLEX

Name of complex	Cellular role	Medical applications
Polyadenylation-complex	RNA processing/modification	infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis; cancer

TABLE 4

CHARACTERIZATION OF PREVIOUSLY UNDESCRIBED PROTEINS

List of proteins of unknown functions	Putative function	Motifs found by sequence analysis	predicted known orthologues
YCL046W		-	
YGR156W		RRM	
YHL035C	multispecific organic anion transporter 1, multidrug resistance-associated protein 2)	3x TM, 2x PFAM: ABC-membrane domain, 2x AAA	no ortholog, but high identity to MRP1-HUMAN (P33527), MRP3-HUMAN (O15438), MRP2-HUMAN (Q92887)
YKL018W	guanine nucleotide binding protein	4x WD40	
YLR221C		-	
YML030W		2x transmembrane	Q9P297 (CLST 11240 protein)
YOR179C	polyadenylation	-	
Legend			
AAA	ATPases associated with a variety of cellular activities		

"missing upon time of publication"

oligos: purchased from MWG, forwards and reverse primers are pre-mixed to a final concentration of 10 micromolar and delivered in a 96 well plate format

step	volume	temperature	Device
prepare master mix (AmpliTaq, Perkin Elmer), containing the TAP cassette vector (1.5 ng/25 microliter reaction)	up to 2.5 ml	4C master mix	cooling platform 1
Dispense the master mix in 96 well plate = reaction plate (96 well plate)	23.5 microliters	4C master mix + reaction plate	cooling platforms 1+ 2
Add 1.0 microliters of oligo (final concentration: 0.4 microM) to reaction plate	1.0 microliters	4C reaction plate	cooling platform 2
Cycle (30 cycles) the PCR reaction			MJ thermocyclers
Pipette 2 microliters from the reaction plate to a 96 well gel (Pharmacia)	2 microliters	rt	Pharmacia ready to run gels
run gel 5 minutes			Pharmacia, Ready-To-Run

3) Yeast transformation

General considerations: Procedure partially automated

Materials: the haploid yeast strain is MGD453-13D: MATa, ade2, arg4, leu2-3,112, trp1-289,ura3-52.

step	volume	temperature	device	"novelties" versus original protocole.
------	--------	-------------	--------	--

The day before, inoculate 200ml of YPD with ~100 microliters of freshly growing yeast culture	200 ml	30C	shaking incubator	
When the culture reaches an OD600 of 1.0, dispense 2 ml of yeast culture in 24 well plate = transformation plate (4)	2ml	rt		uses of 24 well plates (Qiagen)
spin		rt	heraus centrifuge	
aspirate the sup	2 ml	rt		
ressuspend cells in the remaining of media		rt	shaker	no washing steps of the cells
dispense 5 microliters of carrier DNA, (Herring testes Carrier DNA; 10 mg/ml; Clontech)	5 microliters	rt		
mix		rt	shaker	
transfer remaining of PCR from reaction plate to the transformation plate	20 microliters	rt		no cleaning of the PCR reaction
mix		rt	shaker	
add 500 microliters PEG/LiAc	500 microliters (viscous)	rt		
mix		rt	shaker	
add 55 microliters DMSO, mix	50 microliters	rt	shaker	
incubate 15' rt		rt		
incubate 15' at 42C		42C	heating platform with custom made Pelletier	

			elements	
transfer at rt + 700 microliters TE	700 microliters	rt		
spin		rt		
aspirate sup		rt		
add 800 microliters YPD	800 microliters	rt		
incubate 30' at 30 C+shake		30 C	heating platform+shaker	
add 1ml of TE		rt		
spin		rt		
aspirate sup		rt		
add 50 microliters TE	50 microliters	rt	8 tips	
mix		rt	shaker	
plate+ incubate for ~3 days at 30C (10cm Petri dishes or 12-24 wellplates)				

4) Check PCR

general considerations: fully automated. According to results of the transformation 0 to 6 colonies are tested for homologous recombination. These results are filed in an excell file directly linked to the robot program.

Material: the forward oligos are specific for each ORF (for te sequence cf 1); purchased from MWG at 10 micromolar in 96 well plates. The reverse oligo is constant for all ORFs and annealed in the TAP sequence

dispense 20 microliters NaOH (20mM) in 96 well plates	20 microliters	rt	
---	----------------	----	--

pickup colonies (0 to 6) in the 96 well plate containing the NaOH=DNA plate (96 well)			
Boil 2'			
add 140 microliters of TE pH 7.5 (8x dilution)			
spin at 2000 rpm			
Prepare master mix (PCR master, Roche)	up to 9 ml microliters	4C	cooling platform 1
dispense 14.5 microliters in reaction plate (96 well plate)	14.5 microliters	4C	cooling platform 2
add 2.5 microliters of oligo (final concentration 1 microM) to reaction plate	2.5 microliters	4C reaction plate	cooling platform 1
transfer 8 microliters DNA from DNA plate to reaction plate	8 microliters	4C reaction plate	cooling platform 1
Cycle (30 cycles) the PCR reaction			MJ thermocyclers
add 1 microliters of LB	5 microliters	rt	
load 10 microliters on a 96 well agarose gel	10 microliters		Pharmacia ready to run gels
run gel 5 minutes			Pharmacia, Ready-To-Run

5) Dot blot analysis

the remaining of the check PCR positive colonies are restreaked on -ura plates. Plates are incubated at 30C			30C incubator
the next day, the restreak is used to inoculate 2ml of YPD in 24 well plates			30C shaker

(Qiagen). Plates are incubated over night at 30C			
Culture plates are spin			
remove the supernatant			
add 100 microliters of waters			mix shaker
add 100 microliters of 0.2M NaOH			mix shaker
incubate 3 minutes at room temperature			
spin			
remove the supernatant			
add 50 microliters of 2x sample buffer			
boil 3 minutes			
spin			
load on a 96 dot blot apparatus			Biodot, BioRad
dot blot on a nitrocellulose membran			Protran, Schleicher and Schuell
detect TAP tagged protein by ECL using peroxidase anti-peroxidase complex			

TABLE 6

KNOWN COMPONENTS OF THE YEAST mRNA 3'-END PROCESSING MACHINERY

Factor	Function	Poly-peptide comp. (kDa)	Gene product / (ORF)	Sequence motifs
CF IA	Cleavage and polyadenylation	79.8	Rna14	8xHAT domains
		71.9	(YMR061w)	
		50.0	Pcf11 (YDR228c)	
		32.8	Clp1 (YOR250c)	
	Polyadenylation		Rna15 (YGL044c)	RRM
		64.2	Pab1 (YER165w)	
CF IB	Cleavage site selection and polyadenylation	73	Hrp1 (YOL123w)	
CF II/PF I (= CPF)	Cleavage and polyadenylation	153.4	Cft1/Yhh1 (YDR301w)	Lactamase
		96.1	Cft2/Ydh1 (YLR115w)	
		87.6	Ysh1/Brr5 (YLR277c)	NTP_transfer_2
		88.3	Pta1 (YAL043c)	7xWD40
		64.4	Pap1 (YKR002w)	
			Pfs1	5xZnF_C3H1
		53.1	Pfs2 (YNL317w)	
		35.6	Fip1 (YJR093c)	
		24.4	Yth1 (YPR107c)	

NOVEL COMPLEX MEMBERS

Factor	Function	kDA	Gene product (OKE)	Seq. Motifs
--------	----------	-----	--------------------	-------------

	59.7	<u>Ref2 (YDR195w)</u>	---
	59.5	<u>YKL059c = Pfs1?</u>	ZnF_C2HC
	46.8	<u>YGR156w</u>	RRM
	37.0	<u>YKL018w</u>	4xWD40
	35.9	Glc7 (YER133w)	PP2Ac
	23.3	Ssu72 (YNL222w)	---
	20.9	<u>YOR179c</u>	similarity to Ysh1
		Act1	Actin
		Cka1	Kinase
		Eft2	
		Eno2	Enolase
		Gpm1	
		Hhf2	Core Histone
		Hsc82	
		Hta1	Core Histone
		Imd2	
		Imd4	
		Met6	
		Pdc1	
		Pfk1	Phosphofruktokinase
		Sec13	WD domain, G-beta repeat
		Sec31	WD domain, G-beta repeat
		Ssa3	Hsp70 protein domain
		Taf60	---
		Tif4632	Helix-loop-Helix DNA-bind.
		Tsa1	
		Tye7	---
		Vid24	---
		Vps53	---
		YCL046w	ABC transporter transmemb reg
		YHL035c	---
		YLR221c	---
		YML030w	

CF: cleavage factor

PF I: polyadenylation factor

CstF: cleavage and stimulation factor

CPSF: cleavage and polyadenylation specificity factor

YGR156w: has RNA-binding domain

Glc7: was found in Y2H using Ref2 as bait (Uetz screen)

YOR179c: similar to Ysh1 (PF I complex) (37% identity, 56% similarity)

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

Various publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

CLAIMS

1. An isolated complex selected from complex (I) and comprising

(a) a first protein, or a functionally active fragment or functionally active derivative thereof, which first protein is selected from the group consisting of:

(i) Cft1 (SEQ ID NO:3), or a mammalian homolog thereof, or a variant of Cft1 encoded by a nucleic acid that hybridizes to the Cft1 nucleic acid (SEQ ID NO:4) or its complement under low stringency conditions,

(ii) Cft2 (SEQ ID NO:5), or a mammalian homolog thereof, or a variant of Cft2 encoded by a nucleic acid that hybridizes to the Cft2 nucleic acid (SEQ ID NO:6) or its complement under low stringency conditions,

(iii) Clp1 (SEQ ID NO:9), or a mammalian homolog thereof, or a variant of Clp1 encoded by a nucleic acid that hybridizes to the Clp1 nucleic acid (SEQ ID NO:10) or its complement under low stringency conditions,

(iv) Fip1 (SEQ ID NO:19), or a mammalian homolog thereof, or a variant of Fip1 encoded by a nucleic acid that hybridizes to the Fip1 nucleic acid (SEQ ID NO:20) or its complement under low stringency conditions,

(v) Pab1 (SEQ ID NO:33), or a mammalian homolog thereof, or a variant of Pab1 encoded by a nucleic acid that hybridizes to the Pab1 nucleic acid (SEQ ID NO:34) or its complement under low stringency conditions,

(vi) Pap1 (SEQ ID NO:35), or a mammalian homolog thereof, or a variant of Pap1 encoded by a nucleic acid that hybridizes to the Pap1 nucleic acid (SEQ ID NO:36) or its complement under low stringency conditions,

(vii) Pcf11 (SEQ ID NO:37), or a mammalian homolog thereof, or a variant of Pcf11 encoded by a nucleic acid that hybridizes to the Pcf11 nucleic acid (SEQ ID NO:38) or its complement under low stringency conditions,

(viii) Pfs2 (SEQ ID NO:43), or a mammalian homolog thereof, or a variant of Pfs2 encoded by a nucleic acid that hybridizes to the Pfs2 nucleic acid (SEQ ID NO:44) or its complement under low stringency conditions,

(ix) Pta1 (SEQ ID NO:45), or a mammalian homolog thereof, or a variant of Pta1 encoded by a nucleic acid that hybridizes to the Pta1 nucleic acid (SEQ ID NO:46) or its complement under low stringency conditions,

- (x) Rna14 (SEQ ID NO:49), or a mammalian homolog thereof, or a variant of Rna14 encoded by a nucleic acid that hybridizes to the Rna14 nucleic acid (SEQ ID NO:50) or its complement under low stringency conditions,
- (xi) Rna15 (SEQ ID NO:51), or a mammalian homolog thereof, or a variant of Rna15 encoded by a nucleic acid that hybridizes to the Rna15 nucleic acid (SEQ ID NO:52) or its complement under low stringency conditions,
- (xii) Tif4632 (SEQ ID NO:63), or a mammalian homolog thereof, or a variant of Tif4632 encoded by a nucleic acid that hybridizes to the Tif4632 nucleic acid (SEQ ID NO:64) or its complement under low stringency conditions,
- (xiii) Ykl059c (SEQ ID NO:89), or a mammalian homolog thereof, or a variant of Ykl059c encoded by a nucleic acid that hybridizes to the Ykl059c nucleic acid (SEQ ID NO:90) or its complement under low stringency conditions,
- (xiv) Ysh1 (SEQ ID NO:75), or a mammalian homolog thereof, or a variant of Ysh1 encoded by a nucleic acid that hybridizes to the Ysh1 nucleic acid (SEQ ID NO:76) or its complement under low stringency conditions, and
- (xv) Yth1 (SEQ ID NO:77), or a mammalian homolog thereof, or a variant of Yth1 encoded by a nucleic acid that hybridizes to the Yth1 nucleic acid (SEQ ID NO:78) or its complement under low stringency conditions; and
- (b) a second protein, or a functionally active fragment or functionally active derivative thereof, which second protein is selected from the group consisting of:
- (i) Act1 (SEQ ID NO:1), or a mammalian homolog thereof, or a variant of Act1 encoded by a nucleic acid that hybridizes to the Act1 nucleic acid (SEQ ID NO:2) or its complement under low stringency conditions,
- (ii) Cka1 (SEQ ID NO:7), or a mammalian homolog thereof, or a variant of Cka1 encoded by a nucleic acid that hybridizes to the Cka1 nucleic acid (SEQ ID NO:8) or its complement under low stringency conditions,
- (iii) Eft2 (SEQ ID NO:11), or a mammalian homolog thereof, or a variant of Eft2 encoded by a nucleic acid that hybridizes to the Eft2 nucleic acid (SEQ ID NO:12) or its complement under low stringency conditions,
- (iv) Eno2 (SEQ ID NO:13), or a mammalian homolog thereof, or a variant of Eno2 encoded by a nucleic acid that hybridizes to the Eno2 nucleic acid (SEQ ID NO:14) or its complement under low stringency conditions,

(v) Glc7 (SEQ ID NO:15), or a mammalian homolog thereof, or a variant of Glc7 encoded by a nucleic acid that hybridizes to the Glc7 nucleic acid (SEQ ID NO:16) or its complement under low stringency conditions,

(vi) Gpm1 (SEQ ID NO:17), or a mammalian homolog thereof, or a variant of Gpm1 encoded by a nucleic acid that hybridizes to the Gpm1 nucleic acid (SEQ ID NO:18) or its complement under low stringency conditions,

(vii) Hhf2 (SEQ ID NO:21), or a mammalian homolog thereof, or a variant of Hhf2 encoded by a nucleic acid that hybridizes to the Hhf2 nucleic acid (SEQ ID NO:22) or its complement under low stringency conditions,

(viii) Hta1 (SEQ ID NO:23), or a mammalian homolog thereof, or a variant of Hta1 encoded by a nucleic acid that hybridizes to the Hta1 nucleic acid (SEQ ID NO:24) or its complement under low stringency conditions,

(ix) Hsc82 (SEQ ID NO:25), or a mammalian homolog thereof, or a variant of Hsc82 encoded by a nucleic acid that hybridizes to the Hsc82 nucleic acid (SEQ ID NO:26) or its complement under low stringency conditions,

(x) Imd2 (SEQ ID NO:27), or a mammalian homolog thereof, or a variant of Imd2 encoded by a nucleic acid that hybridizes to the Imd2 nucleic acid (SEQ ID NO:28) or its complement under low stringency conditions,

(xi) Imd4 (SEQ ID NO:29), or a mammalian homolog thereof, or a variant of Imd4 encoded by a nucleic acid that hybridizes to the Imd4 nucleic acid (SEQ ID NO:30) or its complement under low stringency conditions,

(xii) Met6 (SEQ ID NO:31), or a mammalian homolog thereof, or a variant of Met6 encoded by a nucleic acid that hybridizes to the Met6 nucleic acid (SEQ ID NO:32) or its complement under low stringency conditions,

(xiii) Pdc1 (SEQ ID NO:39), or a mammalian homolog thereof, or a variant of Pdc1 encoded by a nucleic acid that hybridizes to the Pdc1 nucleic acid (SEQ ID NO:40) or its complement under low stringency conditions,

(xiv) Pfk1 (SEQ ID NO:41), or a mammalian homolog thereof, or a variant of Pfk1 encoded by a nucleic acid that hybridizes to the Pfk1 nucleic acid (SEQ ID NO:42) or its complement under low stringency conditions,

(xv) Ref2 (SEQ ID NO:47), or a mammalian homolog thereof, or a variant of Ref2 encoded by a nucleic acid that hybridizes to the Ref2 nucleic acid (SEQ ID NO:48) or its complement under low stringency conditions,

(xvi) Sec13 (SEQ ID NO:53), or a mammalian homolog thereof, or a variant of Sec13 encoded by a nucleic acid that hybridizes to the Sec13 nucleic acid (SEQ ID NO:54) or its complement under low stringency conditions,

(xvii) Sec31 (SEQ ID NO:55), or a mammalian homolog thereof, or a variant of Sec31 encoded by a nucleic acid that hybridizes to the Sec31 nucleic acid (SEQ ID NO:56) or its complement under low stringency conditions,

(xviii) Ssa3 (SEQ ID NO:57), or a mammalian homolog thereof, or a variant of Ssa3 encoded by a nucleic acid that hybridizes to the Ssa3 nucleic acid (SEQ ID NO:58) or its complement under low stringency conditions,

(xix) Ssu72 (SEQ ID NO:59), or a mammalian homolog thereof, or a variant of Ssu72 encoded by a nucleic acid that hybridizes to the Ssu72 nucleic acid (SEQ ID NO:60) or its complement under low stringency conditions,

(xx) Taf60 (SEQ ID NO:61), or a mammalian homolog thereof, or a variant of Taf60 encoded by a nucleic acid that hybridizes to the Taf60 nucleic acid (SEQ ID NO:62) or its complement under low stringency conditions,

(xxi) Tkl1 (SEQ ID NO:65), or a mammalian homolog thereof, or a variant of Tkl1 encoded by a nucleic acid that hybridizes to the Tkl1 nucleic acid (SEQ ID NO:66) or its complement under low stringency conditions,

(xxii) Tsa1 (SEQ ID NO:67), or a mammalian homolog thereof, or a variant of Tsa1 encoded by a nucleic acid that hybridizes to the Tsa1 nucleic acid (SEQ ID NO:68) or its complement under low stringency conditions,

(xxiii) Tye7 (SEQ ID NO:69), or a mammalian homolog thereof, or a variant of Tye7 encoded by a nucleic acid that hybridizes to the Tye7 nucleic acid (SEQ ID NO:70) or its complement under low stringency conditions,

(xxiv) Vid24 (SEQ ID NO:71), or a mammalian homolog thereof, or a variant of Vid24 encoded by a nucleic acid that hybridizes to the Vid24 nucleic acid (SEQ ID NO:72) or its complement under low stringency conditions,

(xxv) Vps53 (SEQ ID NO:73), or a mammalian homolog thereof, or a variant of Vps53 encoded by a nucleic acid that hybridizes to the Vps53 nucleic acid (SEQ ID NO:74) or its complement under low stringency conditions,

(xxvi) Ycl046w (SEQ ID NO:79), or a mammalian homolog thereof, or a variant of Ycl046w encoded by a nucleic acid that hybridizes to the Ycl046w nucleic acid (SEQ ID NO:80) or its complement under low stringency conditions,

(xxvii) Ygr156w (SEQ ID NO:81), or a mammalian homolog thereof, or a variant of Ygr156w encoded by a nucleic acid that hybridizes to the Ygr156w nucleic acid (SEQ ID NO:82) or its complement under low stringency conditions,

(xxviii) Yhl035c (SEQ ID NO:83), or a mammalian homolog thereof, or a variant of Yhl035c encoded by a nucleic acid that hybridizes to the Yhl035c nucleic acid (SEQ ID NO:84) or its complement under low stringency conditions,

(xxix) Ykl018w (SEQ ID NO:85), or a mammalian homolog thereof, or a variant of Ykl018w encoded by a nucleic acid that hybridizes to the Ykl018w nucleic acid (SEQ ID NO:86) or its complement under low stringency conditions,

(xxx) Ylr221c (SEQ ID NO:87), or a mammalian homolog thereof, or a variant of Ylr221c encoded by a nucleic acid that hybridizes to the Ylr221c nucleic acid (SEQ ID NO:88) or its complement under low stringency conditions,

(xxxi) Yml030w (SEQ ID NO:91), or a mammalian homolog thereof, or a variant of Yml030w encoded by a nucleic acid that hybridizes to the Yml030w nucleic acid (SEQ ID NO:92) or its complement under low stringency conditions, and

(xxxii) Yor179c (SEQ ID NO:93), or a mammalian homolog thereof, or a variant of Yor179c encoded by a nucleic acid that hybridizes to the Yor179c nucleic acid (SEQ ID NO:94) or its complement under low stringency conditions,

wherein said first protein and said second protein are members of a native cellular Polyadenylation-complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C, and a complex (II) comprising at least two second proteins.

2. An isolated complex comprising the following proteins:

(i) Act1 (SEQ ID NO:1), or a mammalian homolog thereof, or a variant of Act1 encoded by a nucleic acid that hybridizes to the Act1 nucleic acid (SEQ ID NO:2) or its complement under low stringency conditions,

(ii) Cft1 (SEQ ID NO:3), or a mammalian homolog thereof, or a variant of Cft1 encoded by a nucleic acid that hybridizes to the Cft1 nucleic acid (SEQ ID NO:4) or its complement under low stringency conditions,

(iii) Cft2 (SEQ ID NO:5), or a mammalian homolog thereof, or a variant of Cft2 encoded by a nucleic acid that hybridizes to the Cft2 nucleic acid (SEQ ID NO:6) or its complement under low stringency conditions,

(iv) Cka1 (SEQ ID NO:7), or a mammalian homolog thereof, or a variant of Cka1 encoded by a nucleic acid that hybridizes to the Cka1 nucleic acid (SEQ ID NO:8) or its complement under low stringency conditions,

(v) Clp1 (SEQ ID NO:9), or a mammalian homolog thereof, or a variant of Clp1 encoded by a nucleic acid that hybridizes to the Clp1 nucleic acid (SEQ ID NO:10) or its complement under low stringency conditions,

(vi) Eft2 (SEQ ID NO:11), or a mammalian homolog thereof, or a variant of Eft2 encoded by a nucleic acid that hybridizes to the Eft2 nucleic acid (SEQ ID NO:12) or its complement under low stringency conditions,

(vii) Eno2 (SEQ ID NO:13), or a mammalian homolog thereof, or a variant of Eno2 encoded by a nucleic acid that hybridizes to the Eno2 nucleic acid (SEQ ID NO:14) or its complement under low stringency conditions,

(viii) Glc7 (SEQ ID NO:15), or a mammalian homolog thereof, or a variant of Glc7 encoded by a nucleic acid that hybridizes to the Glc7 nucleic acid (SEQ ID NO:16) or its complement under low stringency conditions,

(ix) Gpm1 (SEQ ID NO:17), or a mammalian homolog thereof, or a variant of Gpm1 encoded by a nucleic acid that hybridizes to the Gpm1 nucleic acid (SEQ ID NO:18) or its complement under low stringency conditions,

(x) Fip1 (SEQ ID NO:19), or a mammalian homolog thereof, or a variant of Fip1 encoded by a nucleic acid that hybridizes to the Fip1 nucleic acid (SEQ ID NO:20) or its complement under low stringency conditions,

(xi) Hhf2 (SEQ ID NO:21), or a mammalian homolog thereof, or a variant of Hhf2 encoded by a nucleic acid that hybridizes to the Hhf2 nucleic acid (SEQ ID NO:22) or its complement under low stringency conditions,

(xii) Hta1 (SEQ ID NO:23), or a mammalian homolog thereof, or a variant of Hta1 encoded by a nucleic acid that hybridizes to the Hta1 nucleic acid (SEQ ID NO:24) or its complement under low stringency conditions,

(xiii) Hsc82 (SEQ ID NO:25), or a mammalian homolog thereof, or a variant of Hsc82 encoded by a nucleic acid that hybridizes to the Hsc82 nucleic acid (SEQ ID NO:26) or its complement under low stringency conditions,

(xiv) Imd2 (SEQ ID NO:27), or a mammalian homolog thereof, or a variant of Imd2 encoded by a nucleic acid that hybridizes to the Imd2 nucleic acid (SEQ ID NO:28) or its complement under low stringency conditions,

(xv) Imd4 (SEQ ID NO:29), or a mammalian homolog thereof, or a variant of Imd4 encoded by a nucleic acid that hybridizes to the Imd4 nucleic acid (SEQ ID NO:30) or its complement under low stringency conditions,

(xvi) Met6 (SEQ ID NO:31), or a mammalian homolog thereof, or a variant of Met6 encoded by a nucleic acid that hybridizes to the Met6 nucleic acid (SEQ ID NO:32) or its complement under low stringency conditions,

(xvii) Pab1 (SEQ ID NO:33), or a mammalian homolog thereof, or a variant of Pab1 encoded by a nucleic acid that hybridizes to the Pab1 nucleic acid (SEQ ID NO:34) or its complement under low stringency conditions,

(xviii) Pap1 (SEQ ID NO:35), or a mammalian homolog thereof, or a variant of Pap1 encoded by a nucleic acid that hybridizes to the Pap1 nucleic acid (SEQ ID NO:36) or its complement under low stringency conditions,

(xix) Pcf11 (SEQ ID NO:37), or a mammalian homolog thereof, or a variant of Pcf11 encoded by a nucleic acid that hybridizes to the Pcf11 nucleic acid (SEQ ID NO:38) or its complement under low stringency conditions,

(xx) Pdc1 (SEQ ID NO:39), or a mammalian homolog thereof, or a variant of Pdc1 encoded by a nucleic acid that hybridizes to the Pdc1 nucleic acid (SEQ ID NO:40) or its complement under low stringency conditions,

(xxi) Pfk1 (SEQ ID NO:41), or a mammalian homolog thereof, or a variant of Pfk1 encoded by a nucleic acid that hybridizes to the Pfk1 nucleic acid (SEQ ID NO:42) or its complement under low stringency conditions,

(xxii) Pfs2 (SEQ ID NO:43), or a mammalian homolog thereof, or a variant of Pfs2 encoded by a nucleic acid that hybridizes to the Pfs2 nucleic acid (SEQ ID NO:44) or its complement under low stringency conditions,

(xxiii) Pta1 (SEQ ID NO:45), or a mammalian homolog thereof, or a variant of Pta1 encoded by a nucleic acid that hybridizes to the Pta1 nucleic acid (SEQ ID NO:46) or its complement under low stringency conditions,

(xxiv) Ref2 (SEQ ID NO:47), or a mammalian homolog thereof, or a variant of Ref2 encoded by a nucleic acid that hybridizes to the Ref2 nucleic acid (SEQ ID NO:48) or its complement under low stringency conditions,

(xxv) Rna14 (SEQ ID NO:49), or a mammalian homolog thereof, or a variant of Rna14 encoded by a nucleic acid that hybridizes to the Rna14 nucleic acid (SEQ ID NO:50) or its complement under low stringency conditions,

(xxvi) Rna15 (SEQ ID NO:51), or a mammalian homolog thereof, or a variant of Rna15 encoded by a nucleic acid that hybridizes to the Rna15 nucleic acid (SEQ ID NO:52) or its complement under low stringency conditions,

(xxvii) Sec13 (SEQ ID NO:53), or a mammalian homolog thereof, or a variant of Sec13 encoded by a nucleic acid that hybridizes to the Sec13 nucleic acid (SEQ ID NO:54) or its complement under low stringency conditions,

(xxviii) Sec31 (SEQ ID NO:55), or a mammalian homolog thereof, or a variant of Sec31 encoded by a nucleic acid that hybridizes to the Sec31 nucleic acid (SEQ ID NO:56) or its complement under low stringency conditions,

(xxix) Ssa3 (SEQ ID NO:57), or a mammalian homolog thereof, or a variant of Ssa3 encoded by a nucleic acid that hybridizes to the Ssa3 nucleic acid (SEQ ID NO:58) or its complement under low stringency conditions,

(xxx) Ssu72 (SEQ ID NO:59), or a mammalian homolog thereof, or a variant of Ssu72 encoded by a nucleic acid that hybridizes to the Ssu72 nucleic acid (SEQ ID NO:60) or its complement under low stringency conditions,

(xxxi) Taf60 (SEQ ID NO:61), or a mammalian homolog thereof, or a variant of Taf60 encoded by a nucleic acid that hybridizes to the Taf60 nucleic acid (SEQ ID NO:62) or its complement under low stringency conditions,

(xxxii) Tif4632 (SEQ ID NO:63), or a mammalian homolog thereof, or a variant of Tif4632 encoded by a nucleic acid that hybridizes to the Tif4632 nucleic acid (SEQ ID NO:64) or its complement under low stringency conditions,

(xxxiii) Tkl1 (SEQ ID NO:65), or a mammalian homolog thereof, or a variant of Tkl1 encoded by a nucleic acid that hybridizes to the Tkl1 nucleic acid (SEQ ID NO:66) or its complement under low stringency conditions,

(xxxiv) Tsa1 (SEQ ID NO:67), or a mammalian homolog thereof, or a variant of Tsa1 encoded by a nucleic acid that hybridizes to the Tsa1 nucleic acid (SEQ ID NO:68) or its complement under low stringency conditions,

(xxxv) Tye7 (SEQ ID NO:69), or a mammalian homolog thereof, or a variant of Tye7 encoded by a nucleic acid that hybridizes to the Tye7 nucleic acid (SEQ ID NO:70) or its complement under low stringency conditions,

(xxxvi) Vid24 (SEQ ID NO:71), or a mammalian homolog thereof, or a variant of Vid24 encoded by a nucleic acid that hybridizes to the Vid24 nucleic acid (SEQ ID NO:72) or its complement under low stringency conditions,

(xxxvii) Vps53 (SEQ ID NO:73), or a mammalian homolog thereof, or a variant of Vps53 encoded by a nucleic acid that hybridizes to the Vps53 nucleic acid (SEQ ID NO:74) or its complement under low stringency conditions,

(xxxviii) Ysh1 (SEQ ID NO:75), or a mammalian homolog thereof, or a variant of Ysh1 encoded by a nucleic acid that hybridizes to the Ysh1 nucleic acid (SEQ ID NO:76) or its complement under low stringency conditions,

(xxxix) Yth1 (SEQ ID NO:77), or a mammalian homolog thereof, or a variant of Yth1 encoded by a nucleic acid that hybridizes to the Yth1 nucleic acid (SEQ ID NO:78) or its complement under low stringency conditions,

(xl) Ycl046w (SEQ ID NO:79), or a mammalian homolog thereof, or a variant of Ycl046w encoded by a nucleic acid that hybridizes to the Ycl046w nucleic acid (SEQ ID NO:80) or its complement under low stringency conditions,

(xli) Ygr156w (SEQ ID NO:81), or a mammalian homolog thereof, or a variant of Ygr156w encoded by a nucleic acid that hybridizes to the Ygr156w nucleic acid (SEQ ID NO:82) or its complement under low stringency conditions,

(xlii) Yhl035c (SEQ ID NO:83), or a mammalian homolog thereof, or a variant of Yhl035c encoded by a nucleic acid that hybridizes to the Yhl035c nucleic acid (SEQ ID NO:84) or its complement under low stringency conditions,

(xliii) Ykl018w (SEQ ID NO:85), or a mammalian homolog thereof, or a variant of Ykl018w encoded by a nucleic acid that hybridizes to the Ykl018w nucleic acid (SEQ ID NO:86) or its complement under low stringency conditions,

(xliv) Ylr221c (SEQ ID NO:87), or a mammalian homolog thereof, or a variant of Ylr221c encoded by a nucleic acid that hybridizes to the Ylr221c nucleic acid (SEQ ID NO:88) or its complement under low stringency conditions,

(xlv) Ykl059c (SEQ ID NO:89), or a mammalian homolog thereof, or a variant of Ykl059c encoded by a nucleic acid that hybridizes to the Ykl059c nucleic acid (SEQ ID NO:90) or its complement under low stringency conditions,

(xlvii) Yml030w (SEQ ID NO:91), or a mammalian homolog thereof, or a variant of Yml030w encoded by a nucleic acid that hybridizes to the Yml030w nucleic acid (SEQ ID NO:92) or its complement under low stringency conditions, and

(xlviii) Yor179c (SEQ ID NO:93), or a mammalian homolog thereof, or a variant of Yor179c encoded by a nucleic acid that hybridizes to the Yor179c nucleic acid (SEQ ID NO:94) or its complement under low stringency conditions,

wherein said proteins are members of a native cellular Polyadenylation-complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

3. An isolated complex that comprises all but 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27 or 28 of the following 47 proteins:

(i) Act1 (SEQ ID NO:1), or a mammalian homolog thereof, or a variant of Act1 encoded by a nucleic acid that hybridizes to the Act1 nucleic acid (SEQ ID NO:2) or its complement under low stringency conditions,

(ii) Cft1 (SEQ ID NO:3), or a mammalian homolog thereof, or a variant of Cft1 encoded by a nucleic acid that hybridizes to the Cft1 nucleic acid (SEQ ID NO:4) or its complement under low stringency conditions,

(iii) Cft2 (SEQ ID NO:5), or a mammalian homolog thereof, or a variant of Cft2 encoded by a nucleic acid that hybridizes to the Cft2 nucleic acid (SEQ ID NO:6) or its complement under low stringency conditions,

(iv) Cka1 (SEQ ID NO:7), or a mammalian homolog thereof, or a variant of Cka1 encoded by a nucleic acid that hybridizes to the Cka1 nucleic acid (SEQ ID NO:8) or its complement under low stringency conditions,

(v) Clp1 (SEQ ID NO:9), or a mammalian homolog thereof, or a variant of Clp1 encoded by a nucleic acid that hybridizes to the Clp1 nucleic acid (SEQ ID NO:10) or its complement under low stringency conditions,

- (vi) Eft2 (SEQ ID NO:11), or a mammalian homolog thereof, or a variant of Eft2 encoded by a nucleic acid that hybridizes to the Eft2 nucleic acid (SEQ ID NO:12) or its complement under low stringency conditions,
- (vii) Eno2 (SEQ ID NO:13), or a mammalian homolog thereof, or a variant of Eno2 encoded by a nucleic acid that hybridizes to the Eno2 nucleic acid (SEQ ID NO:14) or its complement under low stringency conditions,
- (viii) Glc7 (SEQ ID NO:15), or a mammalian homolog thereof, or a variant of Glc7 encoded by a nucleic acid that hybridizes to the Glc7 nucleic acid (SEQ ID NO:16) or its complement under low stringency conditions,
- (ix) Gpm1 (SEQ ID NO:17), or a mammalian homolog thereof, or a variant of Gpm1 encoded by a nucleic acid that hybridizes to the Gpm1 nucleic acid (SEQ ID NO:18) or its complement under low stringency conditions,
- (x) Fip1 (SEQ ID NO:19), or a mammalian homolog thereof, or a variant of Fip1 encoded by a nucleic acid that hybridizes to the Fip1 nucleic acid (SEQ ID NO:20) or its complement under low stringency conditions,
- (xi) Hhf2 (SEQ ID NO:21), or a mammalian homolog thereof, or a variant of Hhf2 encoded by a nucleic acid that hybridizes to the Hhf2 nucleic acid (SEQ ID NO:22) or its complement under low stringency conditions,
- (xii) Hta1 (SEQ ID NO:23), or a mammalian homolog thereof, or a variant of Hta1 encoded by a nucleic acid that hybridizes to the Hta1 nucleic acid (SEQ ID NO:24) or its complement under low stringency conditions,
- (xiii) Hsc82 (SEQ ID NO:25), or a mammalian homolog thereof, or a variant of Hsc82 encoded by a nucleic acid that hybridizes to the Hsc82 nucleic acid (SEQ ID NO:26) or its complement under low stringency conditions,
- (xiv) Imd2 (SEQ ID NO:27), or a mammalian homolog thereof, or a variant of Imd2 encoded by a nucleic acid that hybridizes to the Imd2 nucleic acid (SEQ ID NO:28) or its complement under low stringency conditions,
- (xv) Imd4 (SEQ ID NO:29), or a mammalian homolog thereof, or a variant of Imd4 encoded by a nucleic acid that hybridizes to the Imd4 nucleic acid (SEQ ID NO:30) or its complement under low stringency conditions,
- (xvi) Met6 (SEQ ID NO:31), or a mammalian homolog thereof, or a variant of Met6 encoded by a nucleic acid that hybridizes to the Met6 nucleic acid (SEQ ID NO:32) or its complement under low stringency conditions,

(xvii) Pab1 (SEQ ID NO:33), or a mammalian homolog thereof, or a variant of Pab1 encoded by a nucleic acid that hybridizes to the Pab1 nucleic acid (SEQ ID NO:34) or its complement under low stringency conditions,

(xviii) Pap1 (SEQ ID NO:35), or a mammalian homolog thereof, or a variant of Pap1 encoded by a nucleic acid that hybridizes to the Pap1 nucleic acid (SEQ ID NO:36) or its complement under low stringency conditions,

(xix) Pcf11 (SEQ ID NO:37), or a mammalian homolog thereof, or a variant of Pcf11 encoded by a nucleic acid that hybridizes to the Pcf11 nucleic acid (SEQ ID NO:38) or its complement under low stringency conditions,

(xx) Pdc1 (SEQ ID NO:39), or a mammalian homolog thereof, or a variant of Pdc1 encoded by a nucleic acid that hybridizes to the Pdc1 nucleic acid (SEQ ID NO:40) or its complement under low stringency conditions,

(xxi) Pfk1 (SEQ ID NO:41), or a mammalian homolog thereof, or a variant of Pfk1 encoded by a nucleic acid that hybridizes to the Pfk1 nucleic acid (SEQ ID NO:42) or its complement under low stringency conditions,

(xxii) Pfs2 (SEQ ID NO:43), or a mammalian homolog thereof, or a variant of Pfs2 encoded by a nucleic acid that hybridizes to the Pfs2 nucleic acid (SEQ ID NO:44) or its complement under low stringency conditions,

(xxiii) Pta1 (SEQ ID NO:45), or a mammalian homolog thereof, or a variant of Pta1 encoded by a nucleic acid that hybridizes to the Pta1 nucleic acid (SEQ ID NO:46) or its complement under low stringency conditions,

(xxiv) Ref2 (SEQ ID NO:47), or a mammalian homolog thereof, or a variant of Ref2 encoded by a nucleic acid that hybridizes to the Ref2 nucleic acid (SEQ ID NO:48) or its complement under low stringency conditions,

(xxv) Rna14 (SEQ ID NO:49), or a mammalian homolog thereof, or a variant of Rna14 encoded by a nucleic acid that hybridizes to the Rna14 nucleic acid (SEQ ID NO:50) or its complement under low stringency conditions,

(xxvi) Rna15 (SEQ ID NO:51), or a mammalian homolog thereof, or a variant of Rna15 encoded by a nucleic acid that hybridizes to the Rna15 nucleic acid (SEQ ID NO:52) or its complement under low stringency conditions,

(xxvii) Sec13 (SEQ ID NO:53), or a mammalian homolog thereof, or a variant of Sec13 encoded by a nucleic acid that hybridizes to the Sec13 nucleic acid (SEQ ID NO:54) or its complement under low stringency conditions,

(xxviii) Sec31 (SEQ ID NO:55), or a mammalian homolog thereof, or a variant of Sec31 encoded by a nucleic acid that hybridizes to the Sec31 nucleic acid (SEQ ID NO:56) or its complement under low stringency conditions,

(xxix) Ssa3 (SEQ ID NO:57), or a mammalian homolog thereof, or a variant of Ssa3 encoded by a nucleic acid that hybridizes to the Ssa3 nucleic acid (SEQ ID NO:58) or its complement under low stringency conditions,

(xxx) Ssu72 (SEQ ID NO:59), or a mammalian homolog thereof, or a variant of Ssu72 encoded by a nucleic acid that hybridizes to the Ssu72 nucleic acid (SEQ ID NO:60) or its complement under low stringency conditions,

(xxxi) Taf60 (SEQ ID NO:61), or a mammalian homolog thereof, or a variant of Taf60 encoded by a nucleic acid that hybridizes to the Taf60 nucleic acid (SEQ ID NO:62) or its complement under low stringency conditions,

(xxxii) Tif4632 (SEQ ID NO:63), or a mammalian homolog thereof, or a variant of Tif4632 encoded by a nucleic acid that hybridizes to the Tif4632 nucleic acid (SEQ ID NO:64) or its complement under low stringency conditions,

(xxxiii) Tkl1 (SEQ ID NO:65), or a mammalian homolog thereof, or a variant of Tkl1 encoded by a nucleic acid that hybridizes to the Tkl1 nucleic acid (SEQ ID NO:66) or its complement under low stringency conditions,

(xxxiv) Tsa1 (SEQ ID NO:67), or a mammalian homolog thereof, or a variant of Tsa1 encoded by a nucleic acid that hybridizes to the Tsa1 nucleic acid (SEQ ID NO:68) or its complement under low stringency conditions,

(xxxv) Tye7 (SEQ ID NO:69), or a mammalian homolog thereof, or a variant of Tye7 encoded by a nucleic acid that hybridizes to the Tye7 nucleic acid (SEQ ID NO:70) or its complement under low stringency conditions,

(xxxvi) Vid24 (SEQ ID NO:71), or a mammalian homolog thereof, or a variant of Vid24 encoded by a nucleic acid that hybridizes to the Vid24 nucleic acid (SEQ ID NO:72) or its complement under low stringency conditions,

(xxxvii) Vps53 (SEQ ID NO:73), or a mammalian homolog thereof, or a variant of Vps53 encoded by a nucleic acid that hybridizes to the Vps53 nucleic acid (SEQ ID NO:74) or its complement under low stringency conditions,

(xxxviii) Ysh1 (SEQ ID NO:75), or a mammalian homolog thereof, or a variant of Ysh1 encoded by a nucleic acid that hybridizes to the Ysh1 nucleic acid (SEQ ID NO:76) or its complement under low stringency conditions,

(xxxix) Yth1 (SEQ ID NO:77), or a mammalian homolog thereof, or a variant of Yth1 encoded by a nucleic acid that hybridizes to the Yth1 nucleic acid (SEQ ID NO:78) or its complement under low stringency conditions,

(xl) Ycl046w (SEQ ID NO:79), or a mammalian homolog thereof, or a variant of Ycl046w encoded by a nucleic acid that hybridizes to the Ycl046w nucleic acid (SEQ ID NO:80) or its complement under low stringency conditions,

(xli) Ygr156w (SEQ ID NO:81), or a mammalian homolog thereof, or a variant of Ygr156w encoded by a nucleic acid that hybridizes to the Ygr156w nucleic acid (SEQ ID NO:82) or its complement under low stringency conditions,

(xlii) Yhl035c (SEQ ID NO:83), or a mammalian homolog thereof, or a variant of Yhl035c encoded by a nucleic acid that hybridizes to the Yhl035c nucleic acid (SEQ ID NO:84) or its complement under low stringency conditions,

(xliii) Ykl018w (SEQ ID NO:85), or a mammalian homolog thereof, or a variant of Ykl018w encoded by a nucleic acid that hybridizes to the Ykl018w nucleic acid (SEQ ID NO:86) or its complement under low stringency conditions,

(xliv) Ylr221c (SEQ ID NO:87), or a mammalian homolog thereof, or a variant of Ylr221c encoded by a nucleic acid that hybridizes to the Ylr221c nucleic acid (SEQ ID NO:88) or its complement under low stringency conditions,

(xlv) Ykl059c (SEQ ID NO:89), or a mammalian homolog thereof, or a variant of Ykl059c encoded by a nucleic acid that hybridizes to the Ykl059c nucleic acid (SEQ ID NO:90) or its complement under low stringency conditions,

(xlvi) Yml030w (SEQ ID NO:91), or a mammalian homolog thereof, or a variant of Yml030w encoded by a nucleic acid that hybridizes to the Yml030w nucleic acid (SEQ ID NO:92) or its complement under low stringency conditions, and

(xlvii) Yor179c (SEQ ID NO:93), or a mammalian homolog thereof, or a variant of Yor179c encoded by a nucleic acid that hybridizes to the Yor179c nucleic acid (SEQ ID NO:94) or its complement under low stringency conditions,

wherein said proteins are members of a native cellular Polyadenylation-complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer

consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

4. The complex according to claim 2, which comprises all but 1 of the 47 proteins.
5. The complex of claim 1, 2, 3 or 4 comprising a functionally active derivative of said first protein and/or a functionally active derivative of said second protein, wherein the functionally active derivative is a fusion protein comprising said first protein or said second protein fused to an amino acid sequence different from the first protein or second protein, respectively.
6. The complex of claim 1, 2, 3 or 4 comprising a fragment of said first protein and/or a fragment of said second protein, which fragment binds to another protein component of said complex.
7. The complex of claim 1, 2, 3, 4, 5 or 6 that is involved in the 3' end processing activity for mRNA.
8. The complex of claim 5 wherein the functionally active derivative is a fusion protein comprising said first protein or said second protein fused to an affinity tag or label.
9. An antibody or a fragment of said antibody containing the binding domain thereof, which binds the complex of claim 1, 2, 3, 4, 5, 6 or 7 and which does not bind the first protein when uncomplexed or the second protein when uncomplexed.
10. A process for processing RNA comprising the step of bringing into contact a product to any of claims 1-8 with RNA, such that the RNA is processed.
11. A pharmaceutical composition comprising the protein complex of claim 1, 2, 3, 4, 5, 6, 7 or 8; and a pharmaceutically acceptable carrier.
12. A method for screening for a molecule that modulates directly or indirectly the function, activity, composition or formation of the complex of any one of claims 1 - 8 comprising the steps of:

(a) exposing said complex, or a cell or organism containing said Polyadenylation-complex to one or more candidate molecules; and

(b) determining the amount of 3' end processing activity for mRNA of, or protein components of, said complex, wherein a change in said amount, activity, or protein components relative to said amount, activity or protein components in the absence of said candidate molecules indicates that the molecules modulate function, activity or composition of said complex.

13. The method of claim 12, wherein the amount of said complex is determined.

14. The method of claim 12, wherein the activity of said complex is determined.

15. The method of claim 14, wherein said determining step comprises isolating from the cell or organism said Polyadenylation-complex to produce said isolated complex and contacting said isolated complex with a RNA molecule such that the complex binds to the RNA.

16. The method of claim 12, wherein the protein components of said complex are determined.

17. The method of claim 16, wherein said determining step comprises determining whether

(i) Act1 (SEQ ID NO:1), or a mammalian homolog thereof, or a variant of Act1 encoded by a nucleic acid that hybridizes to the Act1 nucleic acid (SEQ ID NO:2) or its complement under low stringency conditions,

(ii) Cka1 (SEQ ID NO:7), or a mammalian homolog thereof, or a variant of Cka1 encoded by a nucleic acid that hybridizes to the Cka1 nucleic acid (SEQ ID NO:8) or its complement under low stringency conditions,

(iii) Eft2 (SEQ ID NO:11), or a mammalian homolog thereof, or a variant of Eft2 encoded by a nucleic acid that hybridizes to the Eft2 nucleic acid (SEQ ID NO:12) or its complement under low stringency conditions,

(iv) Eno2 (SEQ ID NO:13), or a mammalian homolog thereof, or a variant of Eno2 encoded by a nucleic acid that hybridizes to the Eno2 nucleic acid (SEQ ID NO:14) or its complement under low stringency conditions,

(v) Glc7 (SEQ ID NO:15), or a mammalian homolog thereof, or a variant of Glc7 encoded by a nucleic acid that hybridizes to the Glc7 nucleic acid (SEQ ID NO:16) or its complement under low stringency conditions,

(vi) Gpm1 (SEQ ID NO:17), or a mammalian homolog thereof, or a variant of Gpm1 encoded by a nucleic acid that hybridizes to the Gpm1 nucleic acid (SEQ ID NO:18) or its complement under low stringency conditions,

(vii) Hhf2 (SEQ ID NO:21), or a mammalian homolog thereof, or a variant of Hhf2 encoded by a nucleic acid that hybridizes to the Hhf2 nucleic acid (SEQ ID NO:22) or its complement under low stringency conditions,

(viii) Hta1 (SEQ ID NO:23), or a mammalian homolog thereof, or a variant of Hta1 encoded by a nucleic acid that hybridizes to the Hta1 nucleic acid (SEQ ID NO:24) or its complement under low stringency conditions,

(ix) Hsc82 (SEQ ID NO:25), or a mammalian homolog thereof, or a variant of Hsc82 encoded by a nucleic acid that hybridizes to the Hsc82 nucleic acid (SEQ ID NO:26) or its complement under low stringency conditions,

(x) Imd2 (SEQ ID NO:27), or a mammalian homolog thereof, or a variant of Imd2 encoded by a nucleic acid that hybridizes to the Imd2 nucleic acid (SEQ ID NO:28) or its complement under low stringency conditions,

(xi) Imd4 (SEQ ID NO:29), or a mammalian homolog thereof, or a variant of Imd4 encoded by a nucleic acid that hybridizes to the Imd4 nucleic acid (SEQ ID NO:30) or its complement under low stringency conditions,

(xii) Met6 (SEQ ID NO:31), or a mammalian homolog thereof, or a variant of Met6 encoded by a nucleic acid that hybridizes to the Met6 nucleic acid (SEQ ID NO:32) or its complement under low stringency conditions,

(xiii) Pdc1 (SEQ ID NO:39), or a mammalian homolog thereof, or a variant of Pdc1 encoded by a nucleic acid that hybridizes to the Pdc1 nucleic acid (SEQ ID NO:40) or its complement under low stringency conditions,

(xiv) Pfk1 (SEQ ID NO:41), or a mammalian homolog thereof, or a variant of Pfk1 encoded by a nucleic acid that hybridizes to the Pfk1 nucleic acid (SEQ ID NO:42) or its complement under low stringency conditions,

(xv) Ref2 (SEQ ID NO:47), or a mammalian homolog thereof, or a variant of Ref2 encoded by a nucleic acid that hybridizes to the Ref2 nucleic acid (SEQ ID NO:48) or its complement under low stringency conditions,

(xvi) Sec13 (SEQ ID NO:53), or a mammalian homolog thereof, or a variant of Sec13 encoded by a nucleic acid that hybridizes to the Sec13 nucleic acid (SEQ ID NO:54) or its complement under low stringency conditions,

(xvii) Sec31 (SEQ ID NO:55), or a mammalian homolog thereof, or a variant of Sec31 encoded by a nucleic acid that hybridizes to the Sec31 nucleic acid (SEQ ID NO:56) or its complement under low stringency conditions,

(xviii) Ssa3 (SEQ ID NO:57), or a mammalian homolog thereof, or a variant of Ssa3 encoded by a nucleic acid that hybridizes to the Ssa3 nucleic acid (SEQ ID NO:58) or its complement under low stringency conditions,

(xix) Ssu72 (SEQ ID NO:59), or a mammalian homolog thereof, or a variant of Ssu72 encoded by a nucleic acid that hybridizes to the Ssu72 nucleic acid (SEQ ID NO:60) or its complement under low stringency conditions,

(xx) Taf60 (SEQ ID NO:61), or a mammalian homolog thereof, or a variant of Taf60 encoded by a nucleic acid that hybridizes to the Taf60 nucleic acid (SEQ ID NO:62) or its complement under low stringency conditions,

(xxi) Tkl1 (SEQ ID NO:65), or a mammalian homolog thereof, or a variant of Tkl1 encoded by a nucleic acid that hybridizes to the Tkl1 nucleic acid (SEQ ID NO:66) or its complement under low stringency conditions,

(xxii) Tsa1 (SEQ ID NO:67), or a mammalian homolog thereof, or a variant of Tsa1 encoded by a nucleic acid that hybridizes to the Tsa1 nucleic acid (SEQ ID NO:68) or its complement under low stringency conditions,

(xxiii) Tye7 (SEQ ID NO:69), or a mammalian homolog thereof, or a variant of Tye7 encoded by a nucleic acid that hybridizes to the Tye7 nucleic acid (SEQ ID NO:70) or its complement under low stringency conditions,

(xxiv) Vid24 (SEQ ID NO:71), or a mammalian homolog thereof, or a variant of Vid24 encoded by a nucleic acid that hybridizes to the Vid24 nucleic acid (SEQ ID NO:72) or its complement under low stringency conditions,

(xxv) Vps53 (SEQ ID NO:73), or a mammalian homolog thereof, or a variant of Vps53 encoded by a nucleic acid that hybridizes to the Vps53 nucleic acid (SEQ ID NO:74) or its complement under low stringency conditions,

(xxvi) Ycl046w (SEQ ID NO:79), or a mammalian homolog thereof, or a variant of Ycl046w encoded by a nucleic acid that hybridizes to the Ycl046w nucleic acid (SEQ ID NO:80) or its complement under low stringency conditions,

(xxvii) Ygr156w (SEQ ID NO:81), or a mammalian homolog thereof, or a variant of Ygr156w encoded by a nucleic acid that hybridizes to the Ygr156w nucleic acid (SEQ ID NO:82) or its complement under low stringency conditions,

(xxviii) Yhl035c (SEQ ID NO:83), or a mammalian homolog thereof, or a variant of Yhl035c encoded by a nucleic acid that hybridizes to the Yhl035c nucleic acid (SEQ ID NO:84) or its complement under low stringency conditions,

(xxix) Ykl018w (SEQ ID NO:85), or a mammalian homolog thereof, or a variant of Ykl018w encoded by a nucleic acid that hybridizes to the Ykl018w nucleic acid (SEQ ID NO:86) or its complement under low stringency conditions,

(xxx) Ylr221c (SEQ ID NO:87), or a mammalian homolog thereof, or a variant of Ylr221c encoded by a nucleic acid that hybridizes to the Ylr221c nucleic acid (SEQ ID NO:88) or its complement under low stringency conditions,

(xxxi) Yml030w (SEQ ID NO:91), or a mammalian homolog thereof, or a variant of Yml030w encoded by a nucleic acid that hybridizes to the Yml030w nucleic acid (SEQ ID NO:92) or its complement under low stringency conditions, and

(xxxii) Yor179c (SEQ ID NO:93), or a mammalian homolog thereof, or a variant of Yor179c encoded by a nucleic acid that hybridizes to the Yor179c nucleic acid (SEQ ID NO:94) or its complement under low stringency conditions, is present in the complex, wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

18. The method of any of claim 12 to 17, wherein said method is a method of screening for a drug for treatment or prevention of a disease or disorder such as infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis; cancer.

19. A method for screening for a molecule that binds the complex of anyone of claim 1 - 8 comprising the following steps:

(a) exposing said complex, or a cell or organism containing said Polyadenylation-complex, to one or more candidate molecules; and

(b) determining whether said complex is bound by any of said candidate molecules.

20. A method for diagnosing or screening for the presence of a disease or disorder or a predisposition for developing a disease or disorder in a subject, which disease or disorder is characterized by an aberrant amount of 3' end processing activity for mRNA of, or component composition of, the complex of any one of the claim 1 - 8, comprising determining the amount of, 3' end processing activity for mRNA of, or protein components of, said complex in a sample derived from a subject, wherein a difference in said amount, activity, or protein components of, said complex in an analogous sample from a subject not having the disease or disorder or predisposition indicates the presence in the subject of the disease or disorder or predisposition.

21. The method of claim 20, wherein the amount of said complex is determined.

22. The method of claim 20, wherein the activity of said complex is determined.

23. The method of claim 22, wherein said determining step comprises isolating from the subject said Polyadenylation-complex to produce said isolated complex and contacting said isolated complex with a RNA molecule such that the complex binds to the RNA.

24. The method of claim 20, wherein the protein components of said complex are determined.

25. The method of claim 24, wherein said determining step comprises determining whether

(i) Act1 (SEQ ID NO:1), or a mammalian homolog thereof, or a variant of Act1 encoded by a nucleic acid that hybridizes to the Act1 nucleic acid (SEQ ID NO:2) or its complement under low stringency conditions,

(ii) Cka1 (SEQ ID NO:7), or a mammalian homolog thereof, or a variant of Cka1 encoded by a nucleic acid that hybridizes to the Cka1 nucleic acid (SEQ ID NO:8) or its complement under low stringency conditions,

(iii) Eft2 (SEQ ID NO:11), or a mammalian homolog thereof, or a variant of Eft2 encoded by a nucleic acid that hybridizes to the Eft2 nucleic acid (SEQ ID NO:12) or its complement under low stringency conditions,

(iv) Eno2 (SEQ ID NO:13), or a mammalian homolog thereof, or a variant of Eno2 encoded by a nucleic acid that hybridizes to the Eno2 nucleic acid (SEQ ID NO:14) or its complement under low stringency conditions,

(v) Glc7 (SEQ ID NO:15), or a mammalian homolog thereof, or a variant of Glc7 encoded by a nucleic acid that hybridizes to the Glc7 nucleic acid (SEQ ID NO:16) or its complement under low stringency conditions,

(vi) Gpm1 (SEQ ID NO:17), or a mammalian homolog thereof, or a variant of Gpm1 encoded by a nucleic acid that hybridizes to the Gpm1 nucleic acid (SEQ ID NO:18) or its complement under low stringency conditions,

(vii) Hhf2 (SEQ ID NO:21), or a mammalian homolog thereof, or a variant of Hhf2 encoded by a nucleic acid that hybridizes to the Hhf2 nucleic acid (SEQ ID NO:22) or its complement under low stringency conditions,

(viii) Hta1 (SEQ ID NO:23), or a mammalian homolog thereof, or a variant of Hta1 encoded by a nucleic acid that hybridizes to the Hta1 nucleic acid (SEQ ID NO:24) or its complement under low stringency conditions,

(ix) Hsc82 (SEQ ID NO:25), or a mammalian homolog thereof, or a variant of Hsc82 encoded by a nucleic acid that hybridizes to the Hsc82 nucleic acid (SEQ ID NO:26) or its complement under low stringency conditions,

(x) Imd2 (SEQ ID NO:27), or a mammalian homolog thereof, or a variant of Imd2 encoded by a nucleic acid that hybridizes to the Imd2 nucleic acid (SEQ ID NO:28) or its complement under low stringency conditions,

(xi) Imd4 (SEQ ID NO:29), or a mammalian homolog thereof, or a variant of Imd4 encoded by a nucleic acid that hybridizes to the Imd4 nucleic acid (SEQ ID NO:30) or its complement under low stringency conditions,

(xii) Met6 (SEQ ID NO:31), or a mammalian homolog thereof, or a variant of Met6 encoded by a nucleic acid that hybridizes to the Met6 nucleic acid (SEQ ID NO:32) or its complement under low stringency conditions,

(xiii) Pdc1 (SEQ ID NO:39), or a mammalian homolog thereof, or a variant of Pdc1 encoded by a nucleic acid that hybridizes to the Pdc1 nucleic acid (SEQ ID NO:40) or its complement under low stringency conditions,

(xiv) Pfk1 (SEQ ID NO:41), or a mammalian homolog thereof, or a variant of Pfk1 encoded by a nucleic acid that hybridizes to the Pfk1 nucleic acid (SEQ ID NO:42) or its complement under low stringency conditions,

(xv) Ref2 (SEQ ID NO:47), or a mammalian homolog thereof, or a variant of Ref2 encoded by a nucleic acid that hybridizes to the Ref2 nucleic acid (SEQ ID NO:48) or its complement under low stringency conditions,

(xvi) Sec13 (SEQ ID NO:53), or a mammalian homolog thereof, or a variant of Sec13 encoded by a nucleic acid that hybridizes to the Sec13 nucleic acid (SEQ ID NO:54) or its complement under low stringency conditions,

(xvii) Sec31 (SEQ ID NO:55), or a mammalian homolog thereof, or a variant of Sec31 encoded by a nucleic acid that hybridizes to the Sec31 nucleic acid (SEQ ID NO:56) or its complement under low stringency conditions,

(xviii) Ssa3 (SEQ ID NO:57), or a mammalian homolog thereof, or a variant of Ssa3 encoded by a nucleic acid that hybridizes to the Ssa3 nucleic acid (SEQ ID NO:58) or its complement under low stringency conditions,

(xix) Ssu72 (SEQ ID NO:59), or a mammalian homolog thereof, or a variant of Ssu72 encoded by a nucleic acid that hybridizes to the Ssu72 nucleic acid (SEQ ID NO:60) or its complement under low stringency conditions,

(xx) Taf60 (SEQ ID NO:61), or a mammalian homolog thereof, or a variant of Taf60 encoded by a nucleic acid that hybridizes to the Taf60 nucleic acid (SEQ ID NO:62) or its complement under low stringency conditions,

(xxi) Tkl1 (SEQ ID NO:65), or a mammalian homolog thereof, or a variant of Tkl1 encoded by a nucleic acid that hybridizes to the Tkl1 nucleic acid (SEQ ID NO:66) or its complement under low stringency conditions,

(xxii) Tsa1 (SEQ ID NO:67), or a mammalian homolog thereof, or a variant of Tsa1 encoded by a nucleic acid that hybridizes to the Tsa1 nucleic acid (SEQ ID NO:68) or its complement under low stringency conditions,

(xxiii) Tye7 (SEQ ID NO:69), or a mammalian homolog thereof, or a variant of Tye7 encoded by a nucleic acid that hybridizes to the Tye7 nucleic acid (SEQ ID NO:70) or its complement under low stringency conditions,

(xxiv) Vid24 (SEQ ID NO:71), or a mammalian homolog thereof, or a variant of Vid24 encoded by a nucleic acid that hybridizes to the Vid24 nucleic acid (SEQ ID NO:72) or its complement under low stringency conditions,

(xxv) Vps53 (SEQ ID NO:73), or a mammalian homolog thereof, or a variant of Vps53 encoded by a nucleic acid that hybridizes to the Vps53 nucleic acid (SEQ ID NO:74) or its complement under low stringency conditions,

(xxvi) Ycl046w (SEQ ID NO:79), or a mammalian homolog thereof, or a variant of Ycl046w encoded by a nucleic acid that hybridizes to the Ycl046w nucleic acid (SEQ ID NO:80) or its complement under low stringency conditions,

(xxvii) Ygr156w (SEQ ID NO:81), or a mammalian homolog thereof, or a variant of Ygr156w encoded by a nucleic acid that hybridizes to the Ygr156w nucleic acid (SEQ ID NO:82) or its complement under low stringency conditions,

(xxviii) Yhl035c (SEQ ID NO:83), or a mammalian homolog thereof, or a variant of Yhl035c encoded by a nucleic acid that hybridizes to the Yhl035c nucleic acid (SEQ ID NO:84) or its complement under low stringency conditions,

(xxix) Ykl018w (SEQ ID NO:85), or a mammalian homolog thereof, or a variant of Ykl018w encoded by a nucleic acid that hybridizes to the Ykl018w nucleic acid (SEQ ID NO:86) or its complement under low stringency conditions,

(xxx) Ylr221c (SEQ ID NO:87), or a mammalian homolog thereof, or a variant of Ylr221c encoded by a nucleic acid that hybridizes to the Ylr221c nucleic acid (SEQ ID NO:88) or its complement under low stringency conditions,

(xxxi) Yml030w (SEQ ID NO:91), or a mammalian homolog thereof, or a variant of Yml030w encoded by a nucleic acid that hybridizes to the Yml030w nucleic acid (SEQ ID NO:92) or its complement under low stringency conditions, and

(xxxii) Yor179c (SEQ ID NO:93), or a mammalian homolog thereof, or a variant of Yor179c encoded by a nucleic acid that hybridizes to the Yor179c nucleic acid (SEQ ID NO:94) or its complement under low stringency conditions, is present in the complex, wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

26. A method for treating or preventing a disease or disorder characterized by an aberrant amount of, 3' end processing activity for mRNA of, or component composition

of, the complex of anyone of claim 1- 8, comprising administering to a subject in need of such treatment or prevention a therapeutically effective amount of one or more molecules that modulate the amount of, 3' end processing activity for mRNA of, or protein components of, said complex.

27. The method according to claim 26, wherein said disease or disorder involves decreased levels of the amount or activity of said complex.

28. The method according to claim 26, wherein said disease or disorder involves increased levels of the amount or activity of said complex.

29. Use of a molecule that modulates the amount of, 3' end processing activity for mRNA of, or the protein components of the complex of any one of claim 1-8 for the manufacture of a medicament for the treatment or prevention of a disease or disorder such as infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis; cancer.

30. A kit comprising in one or more containers

(a) an isolated first protein, or a functionally active fragment or functionally active derivative thereof, which first protein is selected from the group consisting of:

(i) Cft1 (SEQ ID NO:3), or a mammalian homolog thereof, or a variant of Cft1 encoded by a nucleic acid that hybridizes to the Cft1 nucleic acid (SEQ ID NO:4) or its complement under low stringency conditions,

(ii) Cft2 (SEQ ID NO:5), or a mammalian homolog thereof, or a variant of Cft2 encoded by a nucleic acid that hybridizes to the Cft2 nucleic acid (SEQ ID NO:6) or its complement under low stringency conditions,

(iii) Clp1 (SEQ ID NO:9), or a mammalian homolog thereof, or a variant of Clp1 encoded by a nucleic acid that hybridizes to the Clp1 nucleic acid (SEQ ID NO:10) or its complement under low stringency conditions,

(iv) Fip1 (SEQ ID NO:19), or a mammalian homolog thereof, or a variant of Fip1 encoded by a nucleic acid that hybridizes to the Fip1 nucleic acid (SEQ ID NO:20) or its complement under low stringency conditions,

(v) Pab1 (SEQ ID NO:33), or a mammalian homolog thereof, or a variant of Pab1 encoded by a nucleic acid that hybridizes to the Pab1 nucleic acid (SEQ ID NO:34) or its complement under low stringency conditions,

(vi) Pap1 (SEQ ID NO:35), or a mammalian homolog thereof, or a variant of Pap1 encoded by a nucleic acid that hybridizes to the Pap1 nucleic acid (SEQ ID NO:36) or its complement under low stringency conditions,

(vii) Pcf11 (SEQ ID NO:37), or a mammalian homolog thereof, or a variant of Pcf11 encoded by a nucleic acid that hybridizes to the Pcf11 nucleic acid (SEQ ID NO:38) or its complement under low stringency conditions,

(viii) Pfs2 (SEQ ID NO:43), or a mammalian homolog thereof, or a variant of Pfs2 encoded by a nucleic acid that hybridizes to the Pfs2 nucleic acid (SEQ ID NO:44) or its complement under low stringency conditions,

(ix) Pta1 (SEQ ID NO:45), or a mammalian homolog thereof, or a variant of Pta1 encoded by a nucleic acid that hybridizes to the Pta1 nucleic acid (SEQ ID NO:46) or its complement under low stringency conditions,

(x) Rna14 (SEQ ID NO:49), or a mammalian homolog thereof, or a variant of Rna14 encoded by a nucleic acid that hybridizes to the Rna14 nucleic acid (SEQ ID NO:50) or its complement under low stringency conditions,

(xi) Rna15 (SEQ ID NO:51), or a mammalian homolog thereof, or a variant of Rna15 encoded by a nucleic acid that hybridizes to the Rna15 nucleic acid (SEQ ID NO:52) or its complement under low stringency conditions,

(xii) Tif4632 (SEQ ID NO:63), or a mammalian homolog thereof, or a variant of Tif4632 encoded by a nucleic acid that hybridizes to the Tif4632 nucleic acid (SEQ ID NO:64) or its complement under low stringency conditions,

(xiii) Ykl059c (SEQ ID NO:89), or a mammalian homolog thereof, or a variant of Ykl059c encoded by a nucleic acid that hybridizes to the Ykl059c nucleic acid (SEQ ID NO:90) or its complement under low stringency conditions,

(xiv) Ysh1 (SEQ ID NO:75), or a mammalian homolog thereof, or a variant of Ysh1 encoded by a nucleic acid that hybridizes to the Ysh1 nucleic acid (SEQ ID NO:76) or its complement under low stringency conditions, and

(xv) Yth1 (SEQ ID NO:77), or a mammalian homolog thereof, or a variant of Yth1 encoded by a nucleic acid that hybridizes to the Yth1 nucleic acid (SEQ ID NO:78) or its complement under low stringency conditions; and

(b) a second protein, or a functionally active fragment or functionally active derivative thereof, which second protein is selected from the group consisting of:

(i) Act1 (SEQ ID NO:1), or a mammalian homolog thereof, or a variant of Act1 encoded by a nucleic acid that hybridizes to the Act1 nucleic acid (SEQ ID NO:2) or its complement under low stringency conditions,

(ii) Cka1 (SEQ ID NO:7), or a mammalian homolog thereof, or a variant of Cka1 encoded by a nucleic acid that hybridizes to the Cka1 nucleic acid (SEQ ID NO:8) or its complement under low stringency conditions,

(iii) Eft2 (SEQ ID NO:11), or a mammalian homolog thereof, or a variant of Eft2 encoded by a nucleic acid that hybridizes to the Eft2 nucleic acid (SEQ ID NO:12) or its complement under low stringency conditions,

(iv) Eno2 (SEQ ID NO:13), or a mammalian homolog thereof, or a variant of Eno2 encoded by a nucleic acid that hybridizes to the Eno2 nucleic acid (SEQ ID NO:14) or its complement under low stringency conditions,

(v) Glc7 (SEQ ID NO:15), or a mammalian homolog thereof, or a variant of Glc7 encoded by a nucleic acid that hybridizes to the Glc7 nucleic acid (SEQ ID NO:16) or its complement under low stringency conditions,

(vi) Gpm1 (SEQ ID NO:17), or a mammalian homolog thereof, or a variant of Gpm1 encoded by a nucleic acid that hybridizes to the Gpm1 nucleic acid (SEQ ID NO:18) or its complement under low stringency conditions,

(vii) Hhf2 (SEQ ID NO:21), or a mammalian homolog thereof, or a variant of Hhf2 encoded by a nucleic acid that hybridizes to the Hhf2 nucleic acid (SEQ ID NO:22) or its complement under low stringency conditions,

(viii) Hta1 (SEQ ID NO:23), or a mammalian homolog thereof, or a variant of Hta1 encoded by a nucleic acid that hybridizes to the Hta1 nucleic acid (SEQ ID NO:24) or its complement under low stringency conditions,

(ix) Hsc82 (SEQ ID NO:25), or a mammalian homolog thereof, or a variant of Hsc82 encoded by a nucleic acid that hybridizes to the Hsc82 nucleic acid (SEQ ID NO:26) or its complement under low stringency conditions,

(x) Imd2 (SEQ ID NO:27), or a mammalian homolog thereof, or a variant of Imd2 encoded by a nucleic acid that hybridizes to the Imd2 nucleic acid (SEQ ID NO:28) or its complement under low stringency conditions,

(xi) Imd4 (SEQ ID NO:29), or a mammalian homolog thereof, or a variant of Imd4 encoded by a nucleic acid that hybridizes to the Imd4 nucleic acid (SEQ ID NO:30) or its complement under low stringency conditions,

(xii) Met6 (SEQ ID NO:31), or a mammalian homolog thereof, or a variant of Met6 encoded by a nucleic acid that hybridizes to the Met6 nucleic acid (SEQ ID NO:32) or its complement under low stringency conditions,

(xiii) Pdc1 (SEQ ID NO:39), or a mammalian homolog thereof, or a variant of Pdc1 encoded by a nucleic acid that hybridizes to the Pdc1 nucleic acid (SEQ ID NO:40) or its complement under low stringency conditions,

(xiv) Pfk1 (SEQ ID NO:41), or a mammalian homolog thereof, or a variant of Pfk1 encoded by a nucleic acid that hybridizes to the Pfk1 nucleic acid (SEQ ID NO:42) or its complement under low stringency conditions,

(xv) Ref2 (SEQ ID NO:47), or a mammalian homolog thereof, or a variant of Ref2 encoded by a nucleic acid that hybridizes to the Ref2 nucleic acid (SEQ ID NO:48) or its complement under low stringency conditions,

(xvi) Sec13 (SEQ ID NO:53), or a mammalian homolog thereof, or a variant of Sec13 encoded by a nucleic acid that hybridizes to the Sec13 nucleic acid (SEQ ID NO:54) or its complement under low stringency conditions,

(xvii) Sec31 (SEQ ID NO:55), or a mammalian homolog thereof, or a variant of Sec31 encoded by a nucleic acid that hybridizes to the Sec31 nucleic acid (SEQ ID NO:56) or its complement under low stringency conditions,

(xviii) Ssa3 (SEQ ID NO:57), or a mammalian homolog thereof, or a variant of Ssa3 encoded by a nucleic acid that hybridizes to the Ssa3 nucleic acid (SEQ ID NO:58) or its complement under low stringency conditions,

(xix) Ssu72 (SEQ ID NO:59), or a mammalian homolog thereof, or a variant of Ssu72 encoded by a nucleic acid that hybridizes to the Ssu72 nucleic acid (SEQ ID NO:60) or its complement under low stringency conditions,

(xx) Taf60 (SEQ ID NO:61), or a mammalian homolog thereof, or a variant of Taf60 encoded by a nucleic acid that hybridizes to the Taf60 nucleic acid (SEQ ID NO:62) or its complement under low stringency conditions,

(xxi) Tkl1 (SEQ ID NO:65), or a mammalian homolog thereof, or a variant of Tkl1 encoded by a nucleic acid that hybridizes to the Tkl1 nucleic acid (SEQ ID NO:66) or its complement under low stringency conditions,

(xxii) Tsa1 (SEQ ID NO:67), or a mammalian homolog thereof, or a variant of Tsa1 encoded by a nucleic acid that hybridizes to the Tsa1 nucleic acid (SEQ ID NO:68) or its complement under low stringency conditions,

(xxiii) Tye7 (SEQ ID NO:69), or a mammalian homolog thereof, or a variant of Tye7 encoded by a nucleic acid that hybridizes to the Tye7 nucleic acid (SEQ ID NO:70) or its complement under low stringency conditions,

(xxiv) Vid24 (SEQ ID NO:71), or a mammalian homolog thereof, or a variant of Vid24 encoded by a nucleic acid that hybridizes to the Vid24 nucleic acid (SEQ ID NO:72) or its complement under low stringency conditions,

(xxv) Vps53 (SEQ ID NO:73), or a mammalian homolog thereof, or a variant of Vps53 encoded by a nucleic acid that hybridizes to the Vps53 nucleic acid (SEQ ID NO:74) or its complement under low stringency conditions,

(xxvi) Ycl046w (SEQ ID NO:79), or a mammalian homolog thereof, or a variant of Ycl046w encoded by a nucleic acid that hybridizes to the Ycl046w nucleic acid (SEQ ID NO:80) or its complement under low stringency conditions,

(xxvii) Ygr156w (SEQ ID NO:81), or a mammalian homolog thereof, or a variant of Ygr156w encoded by a nucleic acid that hybridizes to the Ygr156w nucleic acid (SEQ ID NO:82) or its complement under low stringency conditions,

(xxviii) Yhl035c (SEQ ID NO:83), or a mammalian homolog thereof, or a variant of Yhl035c encoded by a nucleic acid that hybridizes to the Yhl035c nucleic acid (SEQ ID NO:84) or its complement under low stringency conditions,

(xxix) Ykl018w (SEQ ID NO:85), or a mammalian homolog thereof, or a variant of Ykl018w encoded by a nucleic acid that hybridizes to the Ykl018w nucleic acid (SEQ ID NO:86) or its complement under low stringency conditions,

(xxx) Ylr221c (SEQ ID NO:87), or a mammalian homolog thereof, or a variant of Ylr221c encoded by a nucleic acid that hybridizes to the Ylr221c nucleic acid (SEQ ID NO:88) or its complement under low stringency conditions,

(xxxi) Yml030w (SEQ ID NO:91), or a mammalian homolog thereof, or a variant of Yml030w encoded by a nucleic acid that hybridizes to the Yml030w nucleic acid (SEQ ID NO:92) or its complement under low stringency conditions, and

(xxxii) Yor179c (SEQ ID NO:93), or a mammalian homolog thereof, or a variant of Yor179c encoded by a nucleic acid that hybridizes to the Yor179c nucleic acid (SEQ ID NO:94) or its complement under low stringency conditions,

wherein said first protein and said second protein are members of a native cellular Polyadenylation-complex, and wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 ug/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

31. A kit comprising in a container the isolated complex of any one of claim 1 - 8 or the antibody of claim 9.

32. A kit for processing RNA comprising in a container the isolated complex of any of claims 1-8 optionally together with further components such as reagents and working instructions.

33. A kit for the diagnosis of a disease of mammals, preferentially for a disease or disorder such as infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis or cancer, comprising a product according to any of the claims 1-8 optionally together with further components such as reagents and working instructions.

34. The complex of any one of claim 1 - 8, or the antibody or fragment of claim 9, for use in a method of diagnosing a disease or disorder such as infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis; cancer.

35. A method for the production of a pharmaceutical composition comprising carrying out the method of claim 12 or 19 to identify a molecule that modulates the function, activity, composition or formation of said complex, and further comprising mixing the identified molecule with a pharmaceutically acceptable carrier.

36. A process for preparing complex of claim 1 - 8 and optionally the components thereof comprising the following steps:

- expressing such a protein in a target cell,
- isolating the protein complex which is attached to the tagged protein, and
- optionally disassociating the protein complex and isolating the individual complex members.

37. The process according to claim 36 characterized in that the tagged protein comprises two different tags which allow two separate affinity purification steps.

38. The process according to any of claim 36 - 37 characterized in that two tags are separated by a cleavage site for a protease.

39. Component of the Polyadenylation-complex obtainable by a process according to any of claim 36 - 38.

40. Complex of claim 1 - 8 and/or protein thereof as a target for an active agent of a pharmaceutical, preferably a drug target in the treatment or prevention of a disease or disorder such as infectious diseases; viral infections such as herpes simplex infections, Epstein-Barr-infections, influenza; metabolic disease such as metachromatic leukodystrophy; neurodegenerative disorders such as amyotrophic lateral sclerosis; cancer.

41. Component of the Polyadenylation-complex selected from

a) yeast proteins

- (i) Ycl046w (SEQ ID NO:59),
- (ii) Ygr156w (SEQ ID NO:61),
- (iii) Yhl035c (SEQ ID NO:63),
- (iv) Ykl018w (SEQ ID NO:179),
- (v) Ylr221c (SEQ ID NO:67),
- (vi) Yml030w (SEQ ID NO:69), and
- (vii) Yor179c (SEQ ID NO:71).

b) the mammalian homologs/orthologs of the proteins of (a), and

c) a functionally active fragment or functionally active derivative of the proteins according to (a) and (b) carrying one or more amino acid substitutions, deletions and/or additions.

42. Component as described in claim 41, characterized in that it is encoded by a nucleic acid sequence which hybridizes to a nucleic acid sequence encoding any of the yeast proteins listed in claim 41 under low stringency conditions, wherein said low stringency conditions comprise hybridization in a buffer comprising 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 µg/ml denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20 hours at 40°C, washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1 % SDS for 1.5 hours at 55°C, and washing in a buffer consisting of 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 60°C.

43. Nucleic acid encoding a component to any of claims 41 and 42.

44. Construct, preferably a vector construct, comprising

(a) a nucleic acid according to claim 41 and at least one further nucleic acid which is normally not associated with the nucleic acid according to claim 43,
or

(b) at least two separate nucleic acid sequences each encoding a different protein, or a functionally active fragment or a functionally active derivative thereof at least one of said proteins, or functionally active fragments or functionally active derivative thereof selected from the first group of proteins according to claim 1 (a) and at least one of said proteins, or functionally active fragments or functionally active derivative thereof selected from the second group of proteins according to claim 1 (b).

45. Host cell containing a nucleic acid of claim 43 and/or a construct of claim 44 or containing several vectors comprising on different vectors the nucleic acid sequence encoding at least one of the proteins, or functionally active fragments or functionally active derivatives thereof selected from the first group of proteins according to claim 1(a) and at least one of the proteins, or functionally active fragments or functionally active derivatives thereof selected from the second group of proteins according to claim 1(b).

Fig.1

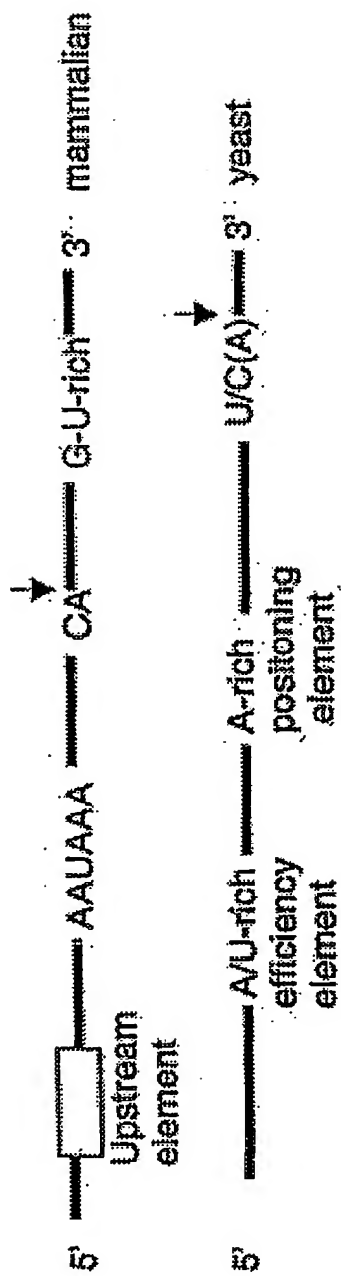


Fig. 2

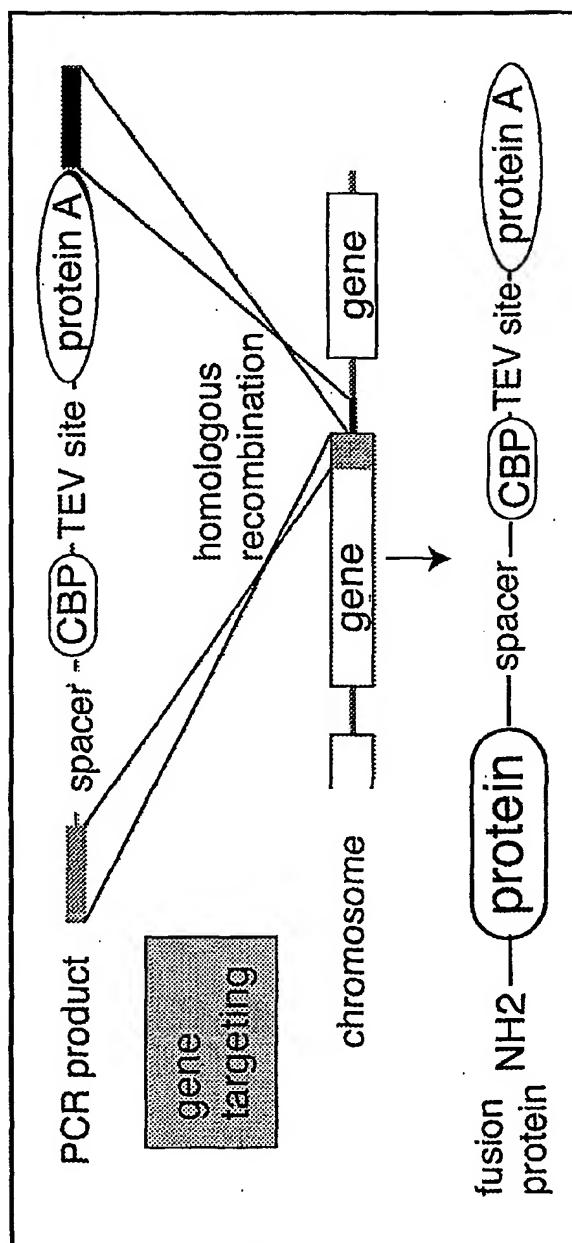
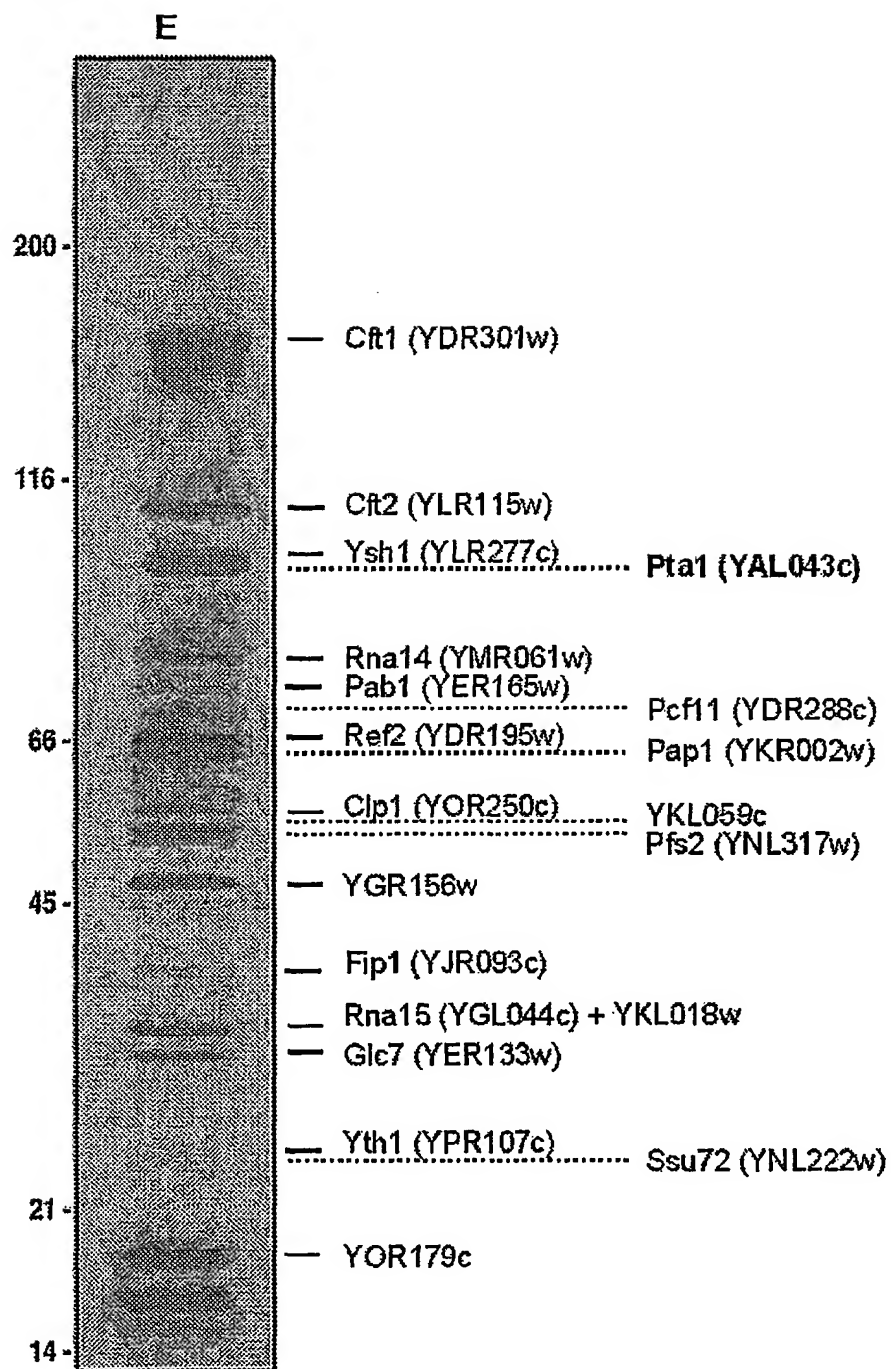
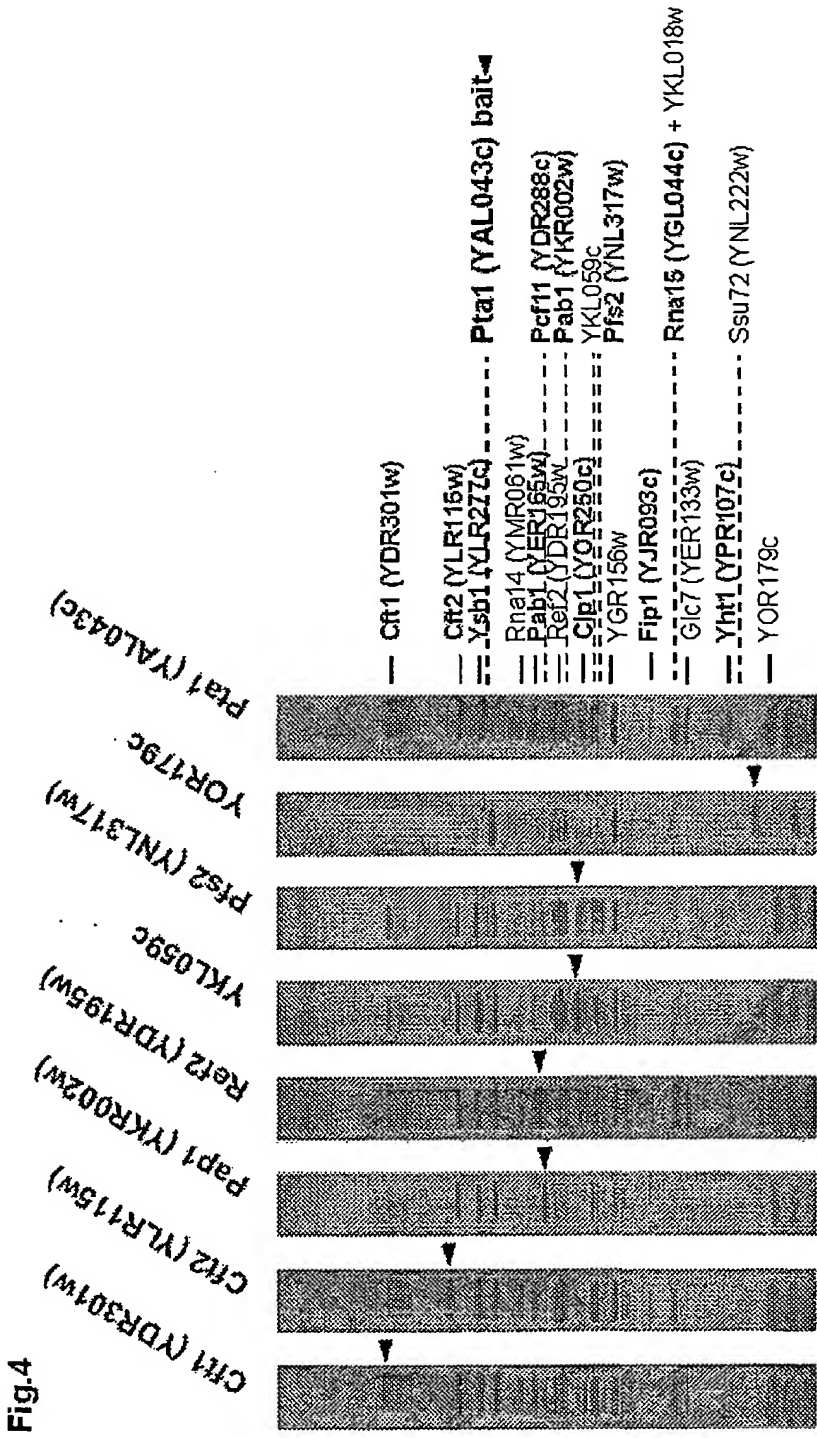


Fig. 3





1/148

SEQUENCE LISTING

<110> Cellzome GmbH

<120> Cleavage and Polyadenylation Complex of Precursor mRNA

<130> PCT1619-031/tc

<140>

<141>

<150> EP 01111774.4

<151> 2001-05-15

<150> EP 01130253.6

<151> 2001-12-20

<160> 94

<170> PatentIn version 3.1

<210> 1

<211> 375

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 1

Met Asp Ser Glu Val Ala Ala Leu Val Ile Asp Asn Gly Ser Gly Met
1 5 10 15

Cys Lys Ala Gly Phe Ala Gly Asp Asp Ala Pro Arg Ala Val Phe Pro
20 25 30

Ser Ile Val Gly Arg Pro Arg His Gln Gly Ile Met Val Gly Met Gly
35 40 45

Gln Lys Asp Ser Tyr Val Gly Asp Glu Ala Gln Ser Lys Arg Gly Ile
50 55 60

Leu Thr Leu Arg Tyr Pro Ile Glu His Gly Ile Val Thr Asn Trp Asp
65 70 75 80

Asp Met Glu Lys Ile Trp His His Thr Phe Tyr Asn Glu Leu Arg Val
85 90 95

Ala Pro Glu Glu His Pro Val Leu Leu Thr Glu Ala Pro Met Asn Pro
100 105 110

Lys Ser Asn Arg Glu Lys Met Thr Gln Ile Met Phe Glu Thr Phe Asn
115 120 125

Val Pro Ala Phe Tyr Val Ser Ile Gln Ala Val Leu Ser Leu Tyr Ser
130 135 140

Ser Gly Arg Thr Thr Gly Ile Val Leu Asp Ser Gly Asp Gly Val Thr
 145 150 155 160

His Val Val Pro Ile Tyr Ala Gly Phe Ser Leu Pro His Ala Ile Leu
 165 170 175

Arg Ile Asp Leu Ala Gly Arg Asp Leu Thr Asp Tyr Leu Met Lys Ile
 180 185 190

Leu Ser Glu Arg Gly Tyr Ser Phe Ser Thr Thr Ala Glu Arg Glu Ile
 195 200 205

Val Arg Asp Ile Lys Glu Lys Leu Cys Tyr Val Ala Leu Asp Phe Glu
 210 215 220

Gln Glu Met Gln Thr Ala Ala Gln Ser Ser Ser Ile Glu Lys Ser Tyr
 225 230 235 240

Glu Leu Pro Asp Gly Gln Val Ile Thr Ile Gly Asn Glu Arg Phe Arg
 245 250 255

Ala Pro Glu Ala Leu Phe His Pro Ser Val Leu Gly Leu Glu Ser Ala
 260 265 270

Gly Ile Asp Gln Thr Thr Tyr Asn Ser Ile Met Lys Cys Asp Val Asp
 275 280 285

Val Arg Lys Glu Leu Tyr Gly Asn Ile Val Met Ser Gly Gly Thr Thr
 290 295 300

Met Phe Pro Gly Ile Ala Glu Arg Met Gln Lys Glu Ile Thr Ala Leu
 305 310 315 320

Ala Pro Ser Ser Met Lys Val Lys Ile Ile Ala Pro Pro Glu Arg Lys
 325 330 335

Tyr Ser Val Trp Ile Gly Gly Ser Ile Leu Ala Ser Leu Thr Thr Phe
 340 345 350

Gln Gln Met Trp Ile Ser Lys Gln Glu Tyr Asp Glu Ser Gly Pro Ser
 355 360 365

Ile Val His His Lys Cys Phe
 370 375

<210> 2

<211> 1125

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 2

```

atggattctg gggttgctgc tttggttatt gataacggtt ctggtatgtg taaagccggt      60
tttgccggtg acgacgctcc tcgtgctgtc ttcccatcta tcgtcggtag accaagacac      120
caaggatatca tggtcgggat gggtcacaaa gactcctacg ttggtgatga agctcaatcc      180
aagagaggta tcttgacttt acgttaccca attgaacacg gtattgtcac caactgggac      240
gatatggaaa agatctggca tcataccttc tacaacgaat tgagagttgc cccagaagaa      300
caccctgttc ttttgactga agctccaatg aaccctaaat caaacagaga aaagatgact      360
caaattatgt ttgaaacttt caacgttcca gccttctacg tttccatcca agccgttttg      420
tccttgactc cttccggtag aactactggg attgttttgg attccgggtga tgggtgttact      480
cacgtcgttc caatttacgc tggtttctct ctacctcacg ccattttgag aatcgatttg      540
gccggtagag atttgactga ctacttgatg aagatcttga gtgaacgtgg ttactctttc      600
tccaccactg ctgaaagaga aattgtccgt gacatcaagg aaaaactatg ttacgtcgcc      660
ttggacttcg aacaagaaat gcaaaccgct gctcaatctt cttcaattga aaaatcctac      720
gaacttccag atggtcaagt catcactatt ggtaacgaaa gattcagagc cccagaagct      780
ttgttccatc cttctgtttt gggtttggaa tctgccggta ttgaccaaac tacttacaac      840
tccatcatga agtgtgatgt cgatgtccgt aaggaattat acggtaacat cgttatgtcc      900
ggtggtacca ccatgttccc aggtattgcc gaaagaatgc aaaaggaaat caccgctttg      960
gctccatctt ccatgaaggt caagatcatt gctcctccag aaagaaagta ctccgtctgg     1020
attggtgggt ctatcttggc ttctttgact accttccaac aaatgtggat ctcaaaacaa     1080
gaatacgacg aaagtgggcc atctatcggt caccacaagt gtttc                               1125

```

<210> 3

<211> 1357

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 3

```

Met Asn Val Tyr Asp Asp Val Leu Asp Ala Thr Val Val Ser His Ser
1              5              10              15

```

```

Leu Ala Thr His Phe Thr Thr Ser Asp Tyr Glu Glu Leu Leu Val Val
                20              25              30

```

```

Arg Thr Asn Ile Leu Ser Val Tyr Arg Pro Thr Arg Asp Gly Lys Leu
          35              40              45

```

```

Tyr Leu Thr Asp Glu Phe Lys Phe His Gly Leu Ile Thr Asp Ile Gly

```

Asn Gly Ala Ile Ile Val Gly Thr Asn Glu Leu Ala Phe Leu Asp Asn
290 295 300

Thr Gly Val Leu Gln Ser Thr Val Leu Leu Asn Ser Phe Ala Asp Lys
 305 310 315 320

Glu Leu Gln Lys Thr Lys Ile Ile Asn Asn Ser Ser Leu Glu Ile Met
 325 330 335

Phe Arg Glu Lys Asn Thr Thr Ser Ile Trp Ile Pro Ser Ser Lys Ser
 340 345 350

Lys Asn Gly Gly Ser Asn Asn Asp Glu Thr Leu Leu Leu Met Asp Leu
 355 360 365

Lys Ser Asn Ile Tyr Tyr Ile Gln Met Glu Ala Glu Gly Arg Leu Leu
 370 375 380

Ile Lys Phe Asp Ile Phe Lys Leu Pro Ile Val Asn Asp Leu Leu Lys
 385 390 395 400

Glu Asn Ser Asn Pro Lys Cys Ile Thr Arg Leu Asn Ala Thr Asn Ser
 405 410 415

Asn Lys Asn Met Asp Leu Phe Ile Gly Phe Gly Ser Gly Asn Ala Leu
 420 425 430

Val Leu Arg Leu Asn Asn Leu Lys Ser Thr Ile Glu Thr Arg Glu Ala
 435 440 445

His Asn Pro Ser Ser Gly Thr Asn Ser Leu Met Asp Ile Asn Asp Asp
 450 455 460

Asp Asp Glu Glu Met Asp Asp Leu Tyr Ala Asp Glu Ala Pro Glu Asn
 465 470 475 480

Gly Leu Thr Thr Asn Asp Ser Lys Gly Thr Val Glu Thr Val Gln Pro
 485 490 495

Phe Asp Ile Glu Leu Leu Ser Ser Leu Arg Asn Val Gly Pro Ile Thr
 500 505 510

Ser Leu Thr Val Gly Lys Val Ser Ser Ile Asp Asp Val Val Lys Gly
 515 520 525

Leu Pro Asn Pro Asn Lys Asn Glu Tyr Ser Leu Val Ala Thr Ser Gly
 530 535 540

Asn Gly Ser Gly Ser His Leu Thr Val Ile Gln Thr Ser Val Gln Pro
 545 550 555 560

Glu Ile Glu Leu Ala Leu Lys Phe Ile Ser Ile Thr Gln Ile Trp Asn
 565 570 575

Leu Lys Ile Lys Gly Arg Asp Arg Tyr Leu Ile Thr Thr Asp Ser Thr
 580 585 590

Lys Ser Arg Ser Asp Ile Tyr Glu Ser Asp Asn Asn Phe Lys Leu His
 595 600 605

Lys Gly Gly Arg Leu Arg Arg Asp Ala Thr Thr Val Tyr Ile Ser Met
 610 615 620

Phe Gly Glu Glu Lys Arg Ile Ile Gln Val Thr Thr Asn His Leu Tyr
 625 630 635 640

Leu Tyr Asp Thr His Phe Arg Arg Leu Thr Thr Ile Lys Phe Asp Tyr
 645 650 655

Glu Val Ile His Val Ser Val Met Asp Pro Tyr Ile Leu Val Thr Val
 660 665 670

Ser Arg Gly Asp Ile Lys Ile Phe Glu Leu Glu Glu Lys Asn Lys Arg
 675 680 685

Lys Leu Leu Lys Val Asp Leu Pro Glu Ile Leu Asn Glu Met Val Ile
 690 695 700

Thr Ser Gly Leu Ile Leu Lys Ser Asn Met Cys Asn Glu Phe Leu Ile
 705 710 715 720

Gly Leu Ser Lys Ser Gln Glu Glu Gln Leu Leu Phe Thr Phe Val Thr
 725 730 735

Ala Asp Asn Gln Ile Ile Phe Phe Thr Lys Asp His Asn Asp Arg Ile
 740 745 750

Phe Gln Leu Asn Gly Val Asp Gln Leu Asn Glu Ser Leu Tyr Ile Ser
 755 760 765

Thr Tyr Gln Leu Gly Asp Glu Ile Val Pro Asp Pro Ser Ile Lys Gln
 770 775 780

Val Met Ile Asn Lys Leu Gly His Asp Asn Lys Glu Glu Tyr Leu Thr
 785 790 795 800

Ile Leu Thr Phe Gly Gly Glu Ile Tyr Gln Tyr Arg Lys Leu Pro Gln
 805 810 815

Arg Arg Ser Arg Phe Tyr Arg Asn Val Thr Arg Asn Asp Leu Ala Ile
 820 825 830

Thr Gly Ala Pro Asp Asn Ala Tyr Ala Lys Gly Val Ser Ser Ile Glu
 835 840 845

Arg Ile Met His Tyr Phe Pro Asp Tyr Asn Gly Tyr Ser Val Ile Phe
 850 855 860

Val Thr Gly Ser Val Pro Tyr Ile Leu Ile Lys Glu Asp Asp Ser Thr
 865 870 875 880

Pro Lys Ile Phe Lys Phe Gly Asn Ile Pro Leu Val Ser Val Thr Pro
 885 890 895

Trp Ser Glu Arg Ser Val Met Cys Val Asp Asp Ile Lys Asn Ala Arg
 900 905 910

Val Tyr Thr Leu Thr Thr Asp Asn Met Tyr Tyr Gly Asn Lys Leu Pro
 915 920 925

Leu Lys Gln Ile Lys Ile Ser Asn Val Leu Asp Asp Tyr Lys Thr Leu
 930 935 940

Gln Lys Leu Val Tyr His Glu Arg Ala Gln Leu Phe Leu Val Ser Tyr
 945 950 955 960

Cys Lys Arg Val Pro Tyr Glu Ala Leu Gly Glu Asp Gly Glu Lys Val
 965 970 975

Ile Gly Tyr Asp Glu Asn Val Pro His Ala Glu Gly Phe Gln Ser Gly
 980 985 990

Ile Leu Leu Ile Asn Pro Lys Ser Trp Lys Val Ile Asp Lys Ile Asp
 995 1000 1005

Phe Pro Lys Asn Ser Val Val Asn Glu Met Arg Ser Ser Met Ile
 1010 1015 1020

Gln Ile Asn Ser Lys Thr Lys Arg Lys Arg Glu Tyr Ile Ile Ala
 1025 1030 1035

Gly Val Ala Asn Ala Thr Thr Glu Asp Thr Pro Pro Thr Gly Ala
 1040 1045 1050

Phe His Ile Tyr Asp Val Ile Glu Val Val Pro Glu Pro Gly Lys

1055		1060		1065
Pro Asp Thr Asn Tyr Lys Leu Lys Glu Ile Phe Gln Glu Glu Val				
1070		1075		1080
Ser Gly Thr Val Ser Thr Val Cys Glu Val Ser Gly Arg Phe Met				
1085		1090		1095
Ile Ser Gln Ser Gln Lys Val Leu Val Arg Asp Ile Gln Glu Asp				
1100		1105		1110
Asn Ser Val Ile Pro Val Ala Phe Leu Asp Ile Pro Val Phe Val				
1115		1120		1125
Thr Asp Ser Lys Ser Phe Gly Asn Leu Leu Ile Ile Gly Asp Ala				
1130		1135		1140
Met Gln Gly Phe Gln Phe Ile Gly Phe Asp Ala Glu Pro Tyr Arg				
1145		1150		1155
Met Ile Ser Leu Gly Arg Ser Met Ser Lys Phe Gln Thr Met Ser				
1160		1165		1170
Leu Glu Phe Leu Val Asn Gly Gly Asp Met Tyr Phe Ala Ala Thr				
1175		1180		1185
Asp Ala Asp Arg Asn Val His Val Leu Lys Tyr Ala Pro Asp Glu				
1190		1195		1200
Pro Asn Ser Leu Ser Gly Gln Arg Leu Val His Cys Ser Ser Phe				
1205		1210		1215
Thr Leu His Ser Thr Asn Ser Cys Met Met Leu Leu Pro Arg Asn				
1220		1225		1230
Glu Glu Phe Gly Ser Pro Gln Val Pro Ser Phe Gln Asn Val Gly				
1235		1240		1245
Gly Gln Val Asp Gly Ser Val Phe Lys Ile Val Pro Leu Ser Glu				
1250		1255		1260
Glu Lys Tyr Arg Arg Leu Tyr Val Ile Gln Gln Gln Ile Ile Asp				
1265		1270		1275
Arg Glu Leu Gln Leu Gly Gly Leu Asn Pro Arg Met Glu Arg Leu				
1280		1285		1290

Ala Asn Asp Phe Tyr Gln Met Gly His Ser Met Arg Pro Met Leu
1295 1300 1305

Asp Phe Asn Val Ile Arg Arg Phe Cys Gly Leu Ala Ile Asp Arg
1310 1315 1320

Arg Lys Ser Ile Ala Gln Lys Ala Gly Arg His Ala His Phe Glu
1325 1330 1335

Ala Trp Arg Asp Ile Ile Asn Ile Glu Phe Ser Met Arg Ser Leu
1340 1345 1350

Cys Gln Gly Lys
1355

<210> 4
<211> 2000
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 4
atgaatgtat atgatgatgt gcttgacgct actggtgtgt ctcattcatt agcaacacat 60
tttactactt cagattatga ggagctctta gtggtgcgaa caaatatcct ttctgtctac 120
aggcctacta gggatggtaa actgtatttg actgatgagt ttaagtttca tggettgate 180
accgatatag gcctcattcc ccaaaaagac agtccgttaa gttgcctgct tttatgcaca 240
ggcgtcgcta aaatctctat tttaaagttc aacacgctaa caaactcaat cgatacccta 300
agcctgcatt actacgaagg aaagttcaag ggtaaatcgt tggtagagct ggccaaaatt 360
tccaccttga gaatggatcc tggaagtctt tgtgccttac tgttcaataa tgacattatc 420
gcttttctac ccttccatgt caataaaaat gacgacgatg aggaggaaga agatgaggac 480
gagaacatag acgatagtga attaatccat agcatgaatc aaaaatccca agggacgaat 540
actttcaata agaggaagag aaccaagtta ggtgacaagt tcacggctcc aagcgtggta 600
ttggtagcga gtgagttgta cgaaggcgcc aaaaatatca tagatattca gttcttgaaa 660
aattttacta aaccaacgat agcgctctta taccaaccaa agcttgtttg ggcaggaaac 720
accaccattt caaaactccc tacacaatac gttatactca cattaaatat ccagcctgca 780
gaaagtgcc acaagattga atcaaccaca atagcatttg tcaaggaact gccttgggat 840
ctacacacca ttgtacctgt ttcgaatggg gctatcattg tgggaaccaa tgaactggca 900
tttctagata atactggcgt tttaaatcg acagtattat taaactcatt tgctgataaa 960
gagttacaga aaacaaaaat tatcaacaat tcttcattgg aaattatgtt caggagaaa 1020
aacactacgt caatttggat accttcgtcc aagagcaaaa acggaggaag taataatgac 1080
gaaacactac tacttatgga cttgaaatcc aatatatact atatccaaat ggaagcagag 1140

```

ggcagggttac tgattaaatt cgatattttc aagctaccta tagtcaatga tcttttgaag 1200
gagaactcca atccaaaatg tataacacgt ttgaatgccca ccaattcaaa caaaaatatg 1260
gacttatttta ttgggttttg ttcaggaaat gctttggttc tcagattgaa taatttgaaa 1320
tctaccattg aaacaagaga agcacacaat ccatcttcag gtacaaatag tttgatggat 1380
atcaatgacg acgatgatga ggaaatggac gatttatacg ccgatgaagc tcctgaaaat 1440
ggactaacia ctaatgattc caaaggcact gtggaaactg ttcaaccttt cgatatcgaa 1500
ctgttatcat ctctaaggaa tgttggtccc atcacatcgt taactgtagg taaggatatct 1560
tccattgatg atgtggtaaa aggactacca aatccaaaca aaaacgagta ttcgctagtt 1620
gccacatcgg gaaatggttc aggttctcat ttgaccgtca tacagactag tgttcaaccg 1680
gaaattgaat tagcattgaa atttatcagt ataacacaaa tctggaatct gaagatcaaa 1740
ggaagagata ggtacttaat aactactgat tccacaaaat cccggagtga catatatgaa 1800
agtgacaata acttcaaact tcataagggc ggccgtttta gaagagatgc tactacagtt 1860
tacatttcaa tgtttggtga agaaaaaga ataatacaag tcaccaccaa tcacttatat 1920
ttatatgata cacattttag acgtctcacc acaattaaat ttgactacga agttattcat 1980
gtttccgtca tggatccgta 2000

```

```

<210> 5
<211> 859
<212> PRT
<213> Saccharomyces cerevisiae

```

```

<400> 5

```

```

Met Thr Tyr Lys Tyr Asn Cys Cys Asp Asp Gly Ser Gly Thr Thr Val
1           5           10           15

```

```

Gly Ser Val Val Arg Phe Asp Asn Val Thr Leu Leu Ile Asp Pro Gly
20           25           30

```

```

Trp Asn Pro Ser Lys Val Ser Tyr Glu Gln Cys Ile Lys Tyr Trp Glu
35           40           45

```

```

Lys Val Ile Pro Glu Ile Asp Val Ile Ile Leu Ser Gln Pro Thr Ile
50           55           60

```

```

Glu Cys Leu Gly Ala His Ser Leu Leu Tyr Tyr Asn Phe Thr Ser His
65           70           75           80

```

```

Phe Ile Ser Arg Ile Gln Val Tyr Ala Thr Leu Pro Val Ile Asn Leu
85           90           95

```

Gly Arg Val Ser Thr Ile Asp Ser Tyr Ala Ser Ala Gly Val Ile Gly
 100 105 110

Pro Tyr Asp Thr Asn Lys Leu Asp Leu Glu Asp Ile Glu Ile Ser Phe
 115 120 125

Asp His Ile Val Pro Leu Lys Tyr Ser Gln Leu Val Asp Leu Arg Ser
 130 135 140

Arg Tyr Asp Gly Leu Thr Leu Leu Ala Tyr Asn Ala Gly Val Cys Pro
 145 150 155 160

Gly Gly Ser Ile Trp Cys Ile Ser Thr Tyr Ser Glu Lys Leu Val Tyr
 165 170 175

Ala Lys Arg Trp Asn His Thr Arg Asp Asn Ile Leu Asn Ala Ala Ser
 180 185 190

Ile Leu Asp Ala Thr Gly Lys Pro Leu Ser Thr Leu Met Arg Pro Ser
 195 200 205

Ala Ile Ile Thr Thr Leu Asp Arg Phe Gly Ser Ser Gln Pro Phe Lys
 210 215 220

Lys Arg Ser Lys Ile Phe Lys Asp Thr Leu Lys Lys Gly Leu Ser Ser
 225 230 235 240

Asp Gly Ser Val Ile Ile Pro Val Asp Met Ser Gly Lys Phe Leu Asp
 245 250 255

Leu Phe Thr Gln Val His Glu Leu Leu Phe Glu Ser Thr Lys Ile Asn
 260 265 270

Ala His Thr Gln Val Pro Val Leu Ile Leu Ser Tyr Ala Arg Gly Arg
 275 280 285

Thr Leu Thr Tyr Ala Lys Ser Met Leu Glu Trp Leu Ser Pro Ser Leu
 290 295 300

Leu Lys Thr Trp Glu Asn Arg Asn Asn Thr Ser Pro Phe Glu Ile Gly
 305 310 315 320

Ser Arg Ile Lys Ile Ile Ala Pro Asn Glu Leu Ser Lys Tyr Pro Gly
 325 330 335

Ser Lys Ile Cys Phe Val Ser Glu Val Gly Ala Leu Ile Asn Glu Val
 340 345 350

Ile Ile Lys Val Gly Asn Ser Glu Lys Thr Thr Leu Ile Leu Thr Lys
 355 360 365

Pro Ser Phe Glu Cys Ala Ser Ser Leu Asp Lys Ile Leu Glu Ile Val
 370 375 380

Glu Gln Asp Glu Arg Asn Trp Lys Thr Phe Pro Glu Asp Gly Lys Ser
 385 390 395 400

Phe Leu Cys Asp Asn Tyr Ile Ser Ile Asp Thr Ile Lys Glu Glu Pro
 405 410 415

Leu Ser Lys Glu Glu Thr Glu Ala Phe Lys Val Gln Leu Lys Glu Lys
 420 425 430

Lys Arg Asp Arg Asn Lys Lys Ile Leu Leu Val Lys Arg Glu Ser Lys
 435 440 445

Lys Leu Ala Asn Gly Asn Ala Ile Ile Asp Asp Thr Asn Gly Glu Arg
 450 455 460

Ala Met Arg Asn Gln Asp Ile Leu Val Glu Asn Val Asn Gly Val Pro
 465 470 475 480

Pro Ile Asp His Ile Met Gly Gly Asp Glu Asp Asp Asp Glu Glu Glu
 485 490 495

Glu Asn Asp Asn Leu Leu Asn Leu Leu Lys Asp Asn Ser Glu Lys Ser
 500 505 510

Ala Ala Lys Lys Asn Thr Glu Val Pro Val Asp Ile Ile Ile Gln Pro
 515 520 525

Ser Ala Ala Ser Lys His Lys Met Phe Pro Phe Asn Pro Ala Lys Ile
 530 535 540

Lys Lys Asp Asp Tyr Gly Thr Val Val Asp Phe Thr Met Phe Leu Pro
 545 550 555 560

Asp Asp Ser Asp Asn Val Asn Gln Asn Ser Arg Lys Arg Pro Leu Lys
 565 570 575

Asp Gly Ala Lys Thr Thr Ser Pro Val Asn Glu Glu Asp Asn Lys Asn
 580 585 590

Glu Glu Glu Asp Gly Tyr Asn Met Ser Asp Pro Ile Ser Lys Arg Ser
 595 600 605

Lys His Arg Ala Ser Arg Tyr Ser Gly Phe Ser Gly Thr Gly Glu Ala
610 615 620

Glu Asn Phe Asp Asn Leu Asp Tyr Leu Lys Ile Asp Lys Thr Leu Ser
625 630 635 640

Lys Arg Thr Ile Ser Thr Val Asn Val Gln Leu Lys Cys Ser Val Val
645 650 655

Ile Leu Asn Leu Gln Ser Leu Val Asp Gln Arg Ser Ala Ser Ile Ile
660 665 670

Trp Pro Ser Leu Lys Ser Arg Lys Ile Val Leu Ser Ala Pro Lys Gln
675 680 685

Ile Gln Asn Glu Glu Ile Thr Ala Lys Leu Ile Lys Lys Asn Ile Glu
690 695 700

Val Val Asn Met Pro Leu Asn Lys Ile Val Glu Phe Ser Thr Thr Ile
705 710 715 720

Lys Thr Leu Asp Ile Ser Ile Asp Ser Asn Leu Asp Asn Leu Leu Lys
725 730 735

Trp Gln Arg Ile Ser Asp Ser Tyr Thr Val Ala Thr Val Val Gly Arg
740 745 750

Leu Val Lys Glu Ser Leu Pro Gln Val Asn Asn His Gln Lys Thr Ala
755 760 765

Ser Arg Ser Lys Leu Val Leu Lys Pro Leu His Gly Ser Ser Arg Ser
770 775 780

His Lys Thr Gly Ala Leu Ser Ile Gly Asp Val Arg Leu Ala Gln Leu
785 790 795 800

Lys Lys Leu Leu Thr Glu Lys Asn Tyr Ile Ala Glu Phe Lys Gly Glu
805 810 815

Gly Thr Leu Val Ile Asn Glu Lys Val Ala Val Arg Lys Ile Asn Asp
820 825 830

Ala Glu Thr Ile Ile Asp Gly Thr Pro Ser Glu Leu Phe Asp Thr Val
835 840 845

Lys Lys Leu Val Thr Asp Met Leu Ala Lys Ile

850

855

<210> 6

<211> 2000

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 6

atgacttata aatacaattg ctgtgatgat ggatcaggaa ccaccgttgg ttcagtggta	60
cgatttgata acgtcacctt gttaatcgat cccggttggga atccatctaa agtatcgat	120
gagcaatgca tcaagtactg ggaaaaagtg ataccggaaa ttgatgtaat aatactatca	180
caaccaacaa ttgaatgttt aggtgctcat tcccttctgt actataattt tacctcgcat	240
ttcatatcca gaattcaagt ttacgcaact cttcctgtta tcaacttagg gagagtttcc	300
accatagatt catatgcac tgcaggtggt ataggtccat atgacacaaa taaactcgat	360
cttgaagata ttgagatata ttttgatcat attgtacctt tgaagtactc gcaattggta	420
gatctacgat caagatacga tggattgact ttactggcgt acaatgctgg tgtgtgtcca	480
ggtggttcta tttggtgcat tagcacatac tcagagaaat tggtttatgc aaagcgatgg	540
aaccacacta gagataacat actaaatgct gcttctatat tggacgccac tggtaagccc	600
ttatcaactt tgatgaggcc ctctgcaata attactactt tggacagggt tggctcctct	660
caaccgttta aaaaacgttc aaaaatcttt aaagatacat taaagaaagg tttgagttcg	720
gatgggtccg ttataatacc ggtagatatg agtggcaaat ttctggacct gttcacacaa	780
gttcatgagt tattatttga aagcacaaag atcaatgcgc acacacaagt acctgtactt	840
attctgtctt acgcgagagg aagaactcta acatatgcca aatcaatgct cgagtggctg	900
tcccttcac tacttaagac atgggaaaac agaaataata cttctccatt tgaaattgga	960
tcacggatca aaatcatcgc accaaacgaa ttgagtaa atcctgggtc aaagatttgc	1020
ttcgtatctg aggtaggagc tttgattaat gaagtataa ttaaagtgg taattctgag	1080
aagacaacct tgattttgac caagccgagc tttgaatgtg cgtcgtcttt agacaagata	1140
ctagaaattg ttgaacaaga tgaaagaaat tggaaaacat ttcctgaaga tggtaaataca	1200
ttcctctgag ataattacat ttcaatagac accataaaag aagaaccctt gagtaaagag	1260
gaaacagagg cgtttaaagt acagcttaa gaaaagaaaa gggacagaaa taagaaaatt	1320
ctactggtga aaagagagtc taagaaactg gctaacggca acgcaatcat agatgatacc	1380
aacggagaga gggcaatgag aaatcaagac attttagtgg aaaacgtcaa tggagtacca	1440
ccgattgatc atattatggg gggtgacgaa gatgatgatg aagaggaaga aaacgacaat	1500
ttgttgaaat tactcaagga taactctgag aaatcagcag cgaaaaaaaa tacggaagtt	1560
cctgtagaca ttataattca accaagtgcc gcttcaaac ataaaatgtt cccatttaac	1620


```

cccgctaaaa tcaagaaaga tgattatggt actggtgtag attttacgat gtttttacct 1680
gatgattctg acaatgttaa tcaaaatagc aggaaaagac ctctcaaaga cggtgctaaa 1740
accacgagcc ctgtaaacga agaagataac aaaaatgaag aagaagatgg ttataatatg 1800
agtgatccaa ttagcaaaaag gagcaaacac cgtgcttctc gttactcggg attttctgga 1860
accggagaag cagaaaattt tgataacctt gactatttaa agatagacaa aacgctctcc 1920
aaaagaacaa tctcaactgt aaatgtccaa ttgaaatgct ccgtagtaat cctaaattta 1980
cagagtttag ttgatcaaag 2000

```

```

<210> 7
<211> 372
<212> PRT
<213> Saccharomyces cerevisiae
<400> 7

```

```

Met Lys Cys Arg Val Trp Ser Glu Ala Arg Val Tyr Thr Asn Ile Asn
1 5 10 15

```

```

Lys Gln Arg Thr Glu Glu Tyr Trp Asp Tyr Glu Asn Thr Val Ile Asp
20 25 30

```

```

Trp Ser Thr Asn Thr Lys Asp Tyr Glu Ile Glu Asn Lys Val Gly Arg
35 40 45

```

```

Gly Lys Tyr Ser Glu Val Phe Gln Gly Val Lys Leu Asp Ser Lys Val
50 55 60

```

```

Lys Ile Val Ile Lys Met Leu Lys Pro Val Lys Lys Lys Lys Ile Lys
65 70 75 80

```

```

Arg Glu Ile Lys Ile Leu Thr Asp Leu Ser Asn Glu Lys Val Pro Pro
85 90 95

```

```

Thr Thr Leu Pro Phe Gln Lys Asp Gln Tyr Tyr Thr Asn Gln Lys Glu
100 105 110

```

```

Asp Val Leu Lys Phe Ile Arg Pro Tyr Ile Phe Asp Gln Pro His Asn
115 120 125

```

```

Gly His Ala Asn Ile Ile His Leu Phe Asp Ile Ile Lys Asp Pro Ile
130 135 140

```

```

Ser Lys Thr Pro Ala Leu Val Phe Glu Tyr Val Asp Asn Val Asp Phe
145 150 155 160

```

```

Arg Ile Leu Tyr Pro Lys Leu Thr Asp Leu Glu Ile Arg Phe Tyr Met

```

	165		170		175
Phe Glu Leu Leu Lys Ala Leu Asp Tyr Cys His Ser Met Gly Ile Met					
	180		185		190
His Arg Asp Val Lys Pro His Asn Val Met Ile Asp His Lys Asn Lys					
	195		200		205
Lys Leu Arg Leu Ile Asp Trp Gly Leu Ala Glu Phe Tyr His Val Asn					
	210		215		220
Met Glu Tyr Asn Val Arg Val Ala Ser Arg Phe Phe Lys Gly Pro Glu					
	225		230		240
Leu Leu Val Asp Tyr Arg Met Tyr Asp Tyr Ser Leu Asp Leu Trp Ser					
	245		250		255
Phe Gly Thr Met Leu Ala Ser Met Ile Phe Lys Arg Glu Pro Phe Phe					
	260		265		270
His Gly Thr Ser Asn Thr Asp Gln Leu Val Lys Ile Val Lys Val Leu					
	275		280		285
Gly Thr Ser Asp Phe Glu Lys Tyr Leu Leu Lys Tyr Glu Ile Thr Leu					
	290		295		300
Pro Arg Glu Phe Tyr Asp Met Asp Gln Tyr Ile Arg Lys Pro Trp His					
	305		310		320
Arg Phe Ile Asn Asp Gly Asn Lys His Leu Ser Gly Asn Asp Glu Ile					
	325		330		335
Ile Asp Leu Ile Asp Asn Leu Leu Arg Tyr Asp His Gln Glu Arg Leu					
	340		345		350
Thr Ala Lys Glu Ala Met Gly His Pro Trp Phe Ala Pro Ile Arg Glu					
	355		360		365
Gln Ile Glu Lys					
	370				

<210> 8

<211> 1119

<212> DNA

<213> Saccharomyces cerevisiae

<400> 8

atgaaatgca gggtatgggtc agaggctcgt gtttatacga atatcaataa acaaagaacc

60

gaggaatatt gggattatga aaatactgta attgattggt ccacaaatac aaaggactat 120
 gaaattgaaa ataaagttgg gcgaggaaaa tactccgagg ttttccaagg tgtcaaatta 180
 gactctaaag ttaaaattgt tattaagatg ttgaaaccag ttaaaaagaa gaagatcaag 240
 agagaaatca aaatcttaac ggatttgtct aacgaaaaag tgcctccaac gactttgcca 300
 tttcaaaaag atcaatatta cacaaatcaa aaggaggacg ttttaaaatt catcaggccc 360
 tatactcttg atcagccaca caatggatcat gcaaacataa ttcattctatt tgatataata 420
 aaggatccca tctcaaaaac tccggctttg gtcttcgaat acgtagataa tgtggacttc 480
 cgtattcttt accctaaatt aaccgatctc gaaattaggt tttatatgtt tgagttatta 540
 aaggccttag actattgtca ttcaatggga ataatgcata gagatgttaa acctcataac 600
 gtaatgattg atcataagaa taaaaaattg cgattgatag attgggggct tgetgaattt 660
 tatcatgtta atatggaata caatgttcgt gttgcgtcga gggtttttaa gggctctgaa 720
 ctactagttg actacagaat gtatgattat tctttagact tgtggctcgt tgggacaatg 780
 ttggcttcta tgatctttaa aagagagcca tttttccatg gaacgagtaa cacagaccag 840
 cttgtcaaga tcgtcaaagt acttggtaca agcgattttg agaaatacct gttgaagtat 900
 gaaattacct taccgagaga attttacgat atggaccaat acatcagaaa gccttggcat 960
 agattcatca atgatggtaa taaacattta agcggcaacg atgaaattat tgaccttatt 1020
 gacaatcttt tgagatatga tcatcaagaa agattaactg ctaaggaagc gatgggacac 1080
 ccgtggtttg cccaataag ggaacaaatt gaaaaataa 1119

<210> 9

<211> 445

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 9

Met Ala Ser Leu Pro Gly Ile Asp Glu His Thr Thr Ser Glu Glu Leu
 1 5 10 15

Ile Thr Gly Asp Asn Glu Trp His Lys Leu Val Ile Pro Lys Gly Ser
 20 25 30

Asp Trp Gln Ile Asp Leu Lys Ala Glu Gly Lys Leu Ile Val Lys Val
 35 40 45

Asn Ser Gly Ile Val Glu Ile Phe Gly Thr Glu Leu Ala Val Asp Asp
 50 55 60

Glu Tyr Thr Phe Gln Asn Trp Lys Phe Pro Ile Tyr Ala Val Glu Glu
 65 70 75 80

Thr Glu Leu Leu Trp Lys Cys Pro Asp Leu Thr Thr Asn Thr Ile Thr
85 90 95

Val Lys Pro Asn His Thr Met Lys Tyr Ile Tyr Asn Leu His Phe Met
100 105 110

Leu Glu Lys Ile Arg Met Ser Asn Phe Glu Gly Pro Arg Val Val Ile
115 120 125

Val Gly Gly Ser Gln Thr Gly Lys Thr Ser Leu Ser Arg Thr Leu Cys
130 135 140

Ser Tyr Ala Leu Lys Phe Asn Ala Tyr Gln Pro Leu Tyr Ile Asn Leu
145 150 155 160

Asp Pro Gln Gln Pro Ile Phe Thr Val Pro Gly Cys Ile Ser Ala Thr
165 170 175

Pro Ile Ser Asp Ile Leu Asp Ala Gln Leu Pro Thr Trp Gly Gln Ser
180 185 190

Leu Thr Ser Gly Ala Thr Leu Leu His Asn Lys Gln Pro Met Val Lys
195 200 205

Asn Phe Gly Leu Glu Arg Ile Asn Glu Asn Lys Asp Leu Tyr Leu Glu
210 215 220

Cys Ile Ser Gln Leu Gly Gln Val Val Gly Gln Arg Leu His Leu Asp
225 230 235 240

Pro Gln Val Arg Arg Ser Gly Cys Ile Val Asp Thr Pro Ser Ile Ser
245 250 255

Gln Leu Asp Glu Asn Leu Ala Glu Leu His His Ile Ile Glu Lys Leu
260 265 270

Asn Val Asn Ile Met Leu Val Leu Cys Ser Glu Thr Asp Pro Leu Trp
275 280 285

Glu Lys Val Lys Lys Thr Phe Gly Pro Glu Leu Gly Asn Asn Asn Ile
290 295 300

Phe Phe Ile Pro Lys Leu Asp Gly Val Ser Ala Val Asp Asp Val Tyr
305 310 315 320

Lys Arg Ser Leu Gln Arg Thr Ser Ile Arg Glu Tyr Phe Tyr Gly Ser
325 330 335

Leu Asp Thr Ala Leu Ser Pro Tyr Ala Ile Gly Val Asp Tyr Glu Asp
 340 345 350

Leu Thr Ile Trp Lys Pro Ser Asn Val Phe Asp Asn Glu Val Gly Arg
 355 360 365

Val Glu Leu Phe Pro Val Thr Ile Thr Pro Ser Asn Leu Gln His Ala
 370 375 380

Ile Ile Ala Ile Thr Phe Ala Glu Arg Arg Ala Asp Gln Ala Thr Val
 385 390 395 400

Ile Lys Ser Pro Ile Leu Gly Phe Ala Leu Ile Thr Glu Val Asn Glu
 405 410 415

Lys Arg Arg Lys Leu Arg Val Leu Leu Pro Val Pro Gly Arg Leu Pro
 420 425 430

Ser Lys Ala Met Ile Leu Thr Ser Tyr Arg Tyr Leu Glu
 435 440 445

<210> 10

<211> 1338

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 10

atggcaagtc tacctggtat tgatgagcat actacctctg aagaattgat aactgggtgat 60
 aacgaatggc acaagttggt tattcccaag ggcagtgatt ggcaaattga tcttaaagca 120
 gagggaaaac tgatagtcaa ggtaaaactcg ggcatagttg agatatttgg caccgagttg 180
 gcagtagatg atgagtacac ttttcagaac tggaagtffc ccatatacgc tgtcgaagaa 240
 acagaattat tatggaaatg tcctgattta actacgaata caataactgt caagcctaac 300
 catactatga aatatattta taatctacac ttcattgttg agaagatacg aatgtctaac 360
 tttgaaggcc ctgagtggt gattgttggt ggttcgcaaa ccgggaagac ttccctatcg 420
 aggacactat gttcttacgc ttgaaattc aacgcttacc aaccactata catcaatctt 480
 gaccctcaac agccattttt tactgttcct ggatgcatat ctgccacccc aatatcagat 540
 atacttgatg cacaactacc cacttggggc cagagtctca ctagtgggtc cacactacta 600
 cataataagc agccaatggt gaaaaatttt ggcctggaaa ggattaatga gaacaaagat 660
 ctataccttg agtgataaag ccagttaggt caagtagtag gtcaaagggt acatttggat 720
 cctcaagtca ggagatcagg gtgcattgtc gatagcccat caatatcaca actggatgaa 780
 aatttggctg aactgcacca tatcattgag aaactcaatg ttaacattat gctagtacta 840

```

tggtctgaaa cggatcctct ttgggagaaa gtaaaaaaga catttggtcc agaattggga    900
aataataata tttttttcat tcctaaattg gatggtgtat ccgcagttga tgatgtatac    960
aaaagatcct tgcagaggac atccataaga gactactttt acggctctct tgatacagcc   1020
ttgagtcctt atgctattgg tgttgattat gaggatttaa ctatttgga acctagcaat   1080
gttttcgata acgaggtcgg tagggtggaa ctgttccccg tcactataac tccaagtaac   1140
ctacagcacg ctattatagc cataacgttc gcagaaagaa gggcagatca ggcaacagta   1200
ataaaatcgc ctattttagg attcgctttg attacagaag ttaatgaaaa aaggcgtaaa   1260
ttaagggttt tactacctgt cccaggccga cttcccagca aggcgatgat tctaacttca   1320
tatagatatt tagagtaa                                     1338

```

```

<210> 11
<211> 842
<212> PRT
<213> Saccharomyces cerevisiae

<400> 11

```

```

Met Val Ala Phe Thr Val Asp Gln Met Arg Ser Leu Met Asp Lys Val
1           5           10          15

```

```

Thr Asn Val Arg Asn Met Ser Val Ile Ala His Val Asp His Gly Lys
          20          25          30

```

```

Ser Thr Leu Thr Asp Ser Leu Val Gln Arg Ala Gly Ile Ile Ser Ala
          35          40          45

```

```

Ala Lys Ala Gly Glu Ala Arg Phe Thr Asp Thr Arg Lys Asp Glu Gln
          50          55          60

```

```

Glu Arg Gly Ile Thr Ile Lys Ser Thr Ala Ile Ser Leu Tyr Ser Glu
65          70          75          80

```

```

Met Ser Asp Glu Asp Val Lys Glu Ile Lys Gln Lys Thr Asp Gly Asn
          85          90          95

```

```

Ser Phe Leu Ile Asn Leu Ile Asp Ser Pro Gly His Val Asp Phe Ser
          100         105         110

```

```

Ser Glu Val Thr Ala Ala Leu Arg Val Thr Asp Gly Ala Leu Val Val
          115         120         125

```

```

Val Asp Thr Ile Glu Gly Val Cys Val Gln Thr Glu Thr Val Leu Arg
          130         135         140

```

Gln Ala Leu Gly Glu Arg Ile Lys Pro Val Val Val Ile Asn Lys Val
 145 150 155 160

Asp Arg Ala Leu Leu Glu Leu Gln Val Ser Lys Glu Asp Leu Tyr Gln
 165 170 175

Thr Phe Ala Arg Thr Val Glu Ser Val Asn Val Ile Val Ser Thr Tyr
 180 185 190

Ala Asp Glu Val Leu Gly Asp Val Gln Val Tyr Pro Ala Arg Gly Thr
 195 200 205

Val Ala Phe Gly Ser Gly Leu His Gly Trp Ala Phe Thr Ile Arg Gln
 210 215 220

Phe Ala Thr Arg Tyr Ala Lys Lys Phe Gly Val Asp Lys Ala Lys Met
 225 230 235 240

Met Asp Arg Leu Trp Gly Asp Ser Phe Phe Asn Pro Lys Thr Lys Lys
 245 250 255

Trp Thr Asn Lys Asp Thr Asp Ala Glu Gly Lys Pro Leu Glu Arg Ala
 260 265 270

Phe Asn Met Phe Ile Leu Asp Pro Ile Phe Arg Leu Phe Thr Ala Ile
 275 280 285

Met Asn Phe Lys Lys Asp Glu Ile Pro Val Leu Leu Glu Lys Leu Glu
 290 295 300

Ile Val Leu Lys Gly Asp Glu Lys Asp Leu Glu Gly Lys Ala Leu Leu
 305 310 315 320

Lys Val Val Met Arg Lys Phe Leu Pro Ala Ala Asp Ala Leu Leu Glu
 325 330 335

Met Ile Val Leu His Leu Pro Ser Pro Val Thr Ala Gln Ala Tyr Arg
 340 345 350

Ala Glu Gln Leu Tyr Glu Gly Pro Ala Asp Asp Ala Asn Cys Ile Ala
 355 360 365

Ile Lys Asn Cys Asp Pro Lys Ala Asp Leu Met Leu Tyr Val Ser Lys
 370 375 380

Met Val Pro Thr Ser Asp Lys Gly Arg Phe Tyr Ala Phe Gly Arg Val
 385 390 395 400

Phe Ala Gly Thr Val Lys Ser Gly Gln Lys Val Arg Ile Gln Gly Pro
 405 410 415

Asn Tyr Val Pro Gly Lys Lys Asp Asp Leu Phe Ile Lys Ala Ile Gln
 420 425 430

Arg Val Val Leu Met Met Gly Arg Phe Val Glu Pro Ile Asp Asp Cys
 435 440 445

Pro Ala Gly Asn Ile Ile Gly Leu Val Gly Ile Asp Gln Phe Leu Leu
 450 455 460

Lys Thr Gly Thr Leu Thr Thr Ser Glu Thr Ala His Asn Met Lys Val
 465 470 475 480

Met Lys Phe Ser Val Ser Pro Val Val Gln Val Ala Val Glu Val Lys
 485 490 495

Asn Ala Asn Asp Leu Pro Lys Leu Val Glu Gly Leu Lys Arg Leu Ser
 500 505 510

Lys Ser Asp Pro Cys Val Leu Thr Tyr Met Ser Glu Ser Gly Glu His
 515 520 525

Ile Val Ala Gly Thr Gly Glu Leu His Leu Glu Ile Cys Leu Gln Asp
 530 535 540

Leu Glu His Asp His Ala Gly Val Pro Leu Lys Ile Ser Pro Pro Val
 545 550 555 560

Val Ala Tyr Arg Glu Thr Val Glu Ser Glu Ser Ser Gln Thr Ala Leu
 565 570 575

Ser Lys Ser Pro Asn Lys His Asn Arg Ile Tyr Leu Lys Ala Glu Pro
 580 585 590

Ile Asp Glu Glu Val Ser Leu Ala Ile Glu Asn Gly Ile Ile Asn Pro
 595 600 605

Arg Asp Asp Phe Lys Ala Arg Ala Arg Ile Met Ala Asp Asp Tyr Gly
 610 615 620

Trp Asp Val Thr Asp Ala Arg Lys Ile Trp Cys Phe Gly Pro Asp Gly
 625 630 635 640

Asn Gly Pro Asn Leu Val Ile Asp Gln Thr Lys Ala Val Gln Tyr Leu
 645 650 655

His Glu Ile Lys Asp Ser Val Val Ala Ala Phe Gln Trp Ala Thr Lys
660 665 670

Glu Gly Pro Ile Phe Gly Glu Glu Met Arg Ser Val Arg Val Asn Ile
675 680 685

Leu Asp Val Thr Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln
690 695 700

Ile Ile Pro Thr Met Arg Arg Ala Thr Tyr Ala Gly Phe Leu Leu Ala
705 710 715 720

Asp Pro Lys Ile Gln Glu Pro Val Phe Leu Val Glu Ile Gln Cys Pro
725 730 735

Glu Gln Ala Val Gly Gly Ile Tyr Ser Val Leu Asn Lys Lys Arg Gly
740 745 750

Gln Val Val Ser Glu Glu Gln Arg Pro Gly Thr Pro Leu Phe Thr Val
755 760 765

Lys Ala Tyr Leu Pro Val Asn Glu Ser Phe Gly Phe Thr Gly Glu Leu
770 775 780

Arg Gln Ala Thr Gly Gly Gln Ala Phe Pro Gln Met Val Phe Asp His
785 790 795 800

Trp Ser Thr Leu Gly Ser Asp Pro Leu Asp Pro Thr Ser Lys Ala Gly
805 810 815

Glu Ile Val Leu Ala Ala Arg Lys Arg His Gly Met Lys Glu Glu Val
820 825 830

Pro Gly Trp Gln Glu Tyr Tyr Asp Lys Leu
835 840

<210> 12

<211> 2529

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 12

atggttgctt tcactgttga ccaaagtcgt tctttaatgg acaaagttac caatgtgcgt 60

aacatgtccg ttattgctca cgtcgatcat ggtaagtcca ctttgaccga ttccttggtc 120

caaagagccg gtattatttc cgtgctaag gctggtgaag ctggtttcac cgataaccaga 180

aaggatgaac aagaaagagg tatcactatc aagtctaccg ctatttctct atactctgaa 240

atgtctgacg aagatgtcaa ggaaatcaag caaaagaccg acggtaaactc cttcttgatc	300
aacttgatcg actctccagg tcacgttgac ttctcctctg aagttactgc cgctttacgt	360
gtcactgacg gtgcttttggg tgcgtcgac accattgaag gtgtctgtgt ccaaaccgaa	420
actgttttga gacaagcttt ggggtgagaga atcaagcctg ttgttggttat caacaaggtc	480
gacagagctt tgttgaatt gcaagtttct aaggaagatt tataccaaac ctttgccaga	540
actgttgaat ccgttaacgt catcgtttcc acctacgccg atgaagtttt ggggtgatgtc	600
caagtttacc cagccagagg taccgttgcc ttccggttccg gtttgccacg ttgggctttc	660
actatccgtc aattcgccac cagatatgtt aagaaattcg gtgtcgacaa ggccaagatg	720
atggacagat tatgggggtga ctctttcttc aacccaaaga ccaagaagtg gaccaacaag	780
gacactgatg ctgaaggtaa gccattggaa agagctttca acatgttcat cttggacca	840
atcttcagat tattcactgc tatcatgaac ttcaagaaag atgaaattcc agttttgcta	900
gaaaagttgg aaattgtctt gaagggtgac gaaaaggact tggaaggtaa ggccttggtg	960
aaggttgtta tgagaaagt cttgccagct gccgatgcct tattggaaat gattgtcttg	1020
cacttgccat ctccagtcac tgctcaagcc tacagagctg aacaattata cgaaggcca	1080
gctgacgatg ccaactgtat tgctatcaag aactgtgatc caaaggctga tttgatgttg	1140
tacgtctcca agatggtgcc aacctctgat aagggtagat tctacgcctt cggtagagtt	1200
tttgccggta ctgttaagtc cgggtcaaaag gtcagaatcc aagggtccaa ctacgttcca	1260
ggtaagaagg acgatttggt catcaaggcc attcaaagag ttgttttgat gatgggtaga	1320
tttgtcgaa caatcgatga ctgtccagcc ggtaacatta tcggtttagt cggtatcgat	1380
caattcttgt tgaagactgg tactttgacc accagtgaag ctgctcaca catgaaggtc	1440
atgaaattct ctgtctctcc agttgtgcaa gtcgctgtcg aagtcaagaa cgctaacgac	1500
ttacaaaaat tggtcgaagg tttgaagaga ttgtccaagt ctgatccatg tgtcttgacc	1560
tatatgtctg aatccggtga acatatcgtt gctggtagcc gtgaattgca tttggaaatt	1620
tgtttgcaag atttgaaca cgaccacgct ggtgttccat tgaagatctc cccaccagtt	1680
gtcgcttaca gagaaactgt tgaaagtga tcttctcaaa ctgctttgtc caagtctcca	1740
aacaagcata acagaatcta cttgaaggct gaaccaattg acgaagaagt ctctttggct	1800
attgaaaacg gtatcatcaa ccaagagat gatttcaagg ccagagctag aatcatggct	1860
gacgactacg gttgggatgt caccgatgcc agaaagatct ggtgtttcgg tccagacggg	1920
aacgggtcaa acttggttat tgaccaaact aaggctgtcc aatacttgca cgaaatcaag	1980
gattccggtt ttgctgcttt ccaatgggct accaaggaag gtccaatttt cgggtgaagaa	2040
atgagatctg tcagagttaa cattttggat gttactttac atgccgatgc tatccacaga	2100

```

ggtaggtgggtc aaatcatccc aaccatgaga agagctactt acgctgggtt cttgttggct 2160
gatccaaaga tccaagaacc agttttcttg gtcgaaattc aatgtccaga acaagccgtc 2220
ggtaggtatct actccgtctt aaacaagaag agaggtcaag tcgtttctga agaacaaga 2280
ccaggtactc cattgtttac cgtcaaggcc tacttgccag ttaacgaatc tttcgggttc 2340
actggtgaat tgagacaagc tactggtggt caagctttcc cacaatggt tttcgaccat 2400
tggtccactt taggttctga cccattggac ccaaccteta aggctggtga aattgttctt 2460
gctgctcgta agagacacgg tatgaaggaa gaagttccag gctggcaaga atattacgac 2520
aaattgtaa 2529

```

<210> 13

<211> 437

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 13

```

Met Ala Val Ser Lys Val Tyr Ala Arg Ser Val Tyr Asp Ser Arg Gly
1             5             10            15

```

```

Asn Pro Thr Val Glu Val Glu Leu Thr Thr Glu Lys Gly Val Phe Arg
                20             25            30

```

```

Ser Ile Val Pro Ser Gly Ala Ser Thr Gly Val His Glu Ala Leu Glu
          35             40            45

```

```

Met Arg Asp Glu Asp Lys Ser Lys Trp Met Gly Lys Gly Val Met Asn
          50             55            60

```

```

Ala Val Asn Asn Val Asn Asn Val Ile Ala Ala Ala Phe Val Lys Ala
65             70             75            80

```

```

Asn Leu Asp Val Lys Asp Gln Lys Ala Val Asp Asp Phe Leu Leu Ser
          85             90            95

```

```

Leu Asp Gly Thr Ala Asn Lys Ser Lys Leu Gly Ala Asn Ala Ile Leu
          100            105           110

```

```

Gly Val Ser Met Ala Ala Ala Arg Ala Ala Ala Ala Glu Lys Asn Val
          115            120           125

```

```

Pro Leu Tyr Gln His Leu Ala Asp Leu Ser Lys Ser Lys Thr Ser Pro
          130            135           140

```

```

Tyr Val Leu Pro Val Pro Phe Leu Asn Val Leu Asn Gly Gly Ser His
145            150            155           160

```

Ala Gly Gly Ala Leu Ala Leu Gln Glu Phe Met Ile Ala Pro Thr Gly
 165 170 175

Ala Lys Thr Phe Ala Glu Ala Met Arg Ile Gly Ser Glu Val Tyr His
 180 185 190

Asn Leu Lys Ser Leu Thr Lys Lys Arg Tyr Gly Ala Ser Ala Gly Asn
 195 200 205

Val Gly Asp Glu Gly Gly Val Ala Pro Asn Ile Gln Thr Ala Glu Glu
 210 215 220

Ala Leu Asp Leu Ile Val Asp Ala Ile Lys Ala Ala Gly His Asp Gly
 225 230 235 240

Lys Val Lys Ile Gly Leu Asp Cys Ala Ser Ser Glu Phe Phe Lys Asp
 245 250 255

Gly Lys Tyr Asp Leu Asp Phe Lys Asn Pro Glu Ser Asp Lys Ser Lys
 260 265 270

Trp Leu Thr Gly Val Glu Leu Ala Asp Met Tyr His Ser Leu Met Lys
 275 280 285

Arg Tyr Pro Ile Val Ser Ile Glu Asp Pro Phe Ala Glu Asp Asp Trp
 290 295 300

Glu Ala Trp Ser His Phe Phe Lys Thr Ala Gly Ile Gln Ile Val Ala
 305 310 315 320

Asp Asp Leu Thr Val Thr Asn Pro Ala Arg Ile Ala Thr Ala Ile Glu
 325 330 335

Lys Lys Ala Ala Asp Ala Leu Leu Leu Lys Val Asn Gln Ile Gly Thr
 340 345 350

Leu Ser Glu Ser Ile Lys Ala Ala Gln Asp Ser Phe Ala Ala Asn Trp
 355 360 365

Gly Val Met Val Ser His Arg Ser Gly Glu Thr Glu Asp Thr Phe Ile
 370 375 380

Ala Asp Leu Val Val Gly Leu Arg Thr Gly Gln Ile Lys Thr Gly Ala
 385 390 395 400

Pro Ala Arg Ser Glu Arg Leu Ala Lys Leu Asn Gln Leu Leu Arg Ile
 405 410 415

Glu Glu Glu Leu Gly Asp Lys Ala Val Tyr Ala Gly Glu Asn Phe His
 420 425 430

His Gly Asp Lys Leu
 435

<210> 14
 <211> 1314
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 14
 atggctgtct ctaaagttta cgctagatcc gtctacgact cccgtggtaa cccaaccgtc 60
 gaagtccaat taaccaccga aaaggggtgtt ttcagatcca ttgttccatc tgggtgcctcc 120
 accgggtgtcc acgaagcttt ggaaatgaga gatgaagaca aatccaagtg gatgggtaag 180
 ggtgttatga acgctgtcaa caacgtcaac aacgtcattg ctgctgcttt cgtcaaggcc 240
 aacctagatg ttaaggacca aaaggccgtc gatgacttct tgttgtcttt ggatgggtacc 300
 gccacaagt ccaagttggg tgctaacgct atcttgggtg tctccatggc cgctgctaga 360
 gccgctgctg ctgaaaagaa cgtcccattg taccaacatt tggtgactt gtctaagtcc 420
 aagacctctc catacgtttt gccagttcca ttcttgaacg ttttgaacgg tggttcccac 480
 gctgggtgtg ctttggcttt gcaagaattc atgattgtc caactgggtg taagaccttc 540
 gctgaagcca tgagaattgg ttccgaagtt taccacaact tgaagtcttt gaccaagaag 600
 agatacggtg cttctgccgg taacgtcggg gacgaagggt gtgttgctcc aacattcaa 660
 accgctgaag aagcttttga cttgattgtt gacgctatca aggtgctgg tcacgacggg 720
 aaggtaaga tcggtttga ctgtgcttcc tctgaattct tcaaggacgg taagtacgac 780
 ttggacttca agaaccaga atctgacaaa tccaagtggg tgactgggtg cgaattagct 840
 gacatgtacc actccttgat gaagagatac ccaattgtct ccatcgaaga tccatttgct 900
 gaagatgact ggggaagcttg gtctcacttc tcaagaccg ctggtatcca aattgttgct 960
 gatgacttga ctgtcaccaa cccagctaga attgctaccg ccatcgaaaa gaaggctgct 1020
 gacgctttgt tgttgaagg taaccaaatac ggtaccttgt ctgaatccat caaggctgct 1080
 caagactctt tcgctgccaa ctgggggtgtt atgggttccc acagatctgg tgaaactgaa 1140
 gacactttca ttgctgactt ggttgctcgg ttgagaactg gtcaaataca gactgggtgct 1200
 ccagctagat ccgaaagatt ggctaagttg aaccaattgt tgagaatcga agaagaattg 1260
 ggtgacaagg ctgtctacgc cggtgaaaac ttccaccacg gtgacaagtt gtaa 1314

<210> 15
 <211> 312

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 15

Met Asp Ser Gln Pro Val Asp Val Asp Asn Ile Ile Asp Arg Leu Leu
 1 5 10 15

Glu Val Arg Gly Ser Lys Pro Gly Gln Gln Val Asp Leu Glu Glu Asn
 20 25 30

Glu Ile Arg Tyr Leu Cys Ser Lys Ala Arg Ser Ile Phe Ile Lys Gln
 35 40 45

Pro Ile Leu Leu Glu Leu Glu Ala Pro Ile Lys Ile Cys Gly Asp Ile
 50 55 60

His Gly Gln Tyr Tyr Asp Leu Leu Arg Leu Phe Glu Tyr Gly Gly Phe
 65 70 75 80

Pro Pro Glu Ser Asn Tyr Leu Phe Leu Gly Asp Tyr Val Asp Arg Gly
 85 90 95

Lys Gln Ser Leu Glu Thr Ile Cys Leu Leu Leu Ala Tyr Lys Ile Lys
 100 105 110

Tyr Pro Glu Asn Phe Phe Ile Leu Arg Gly Asn His Glu Cys Ala Ser
 115 120 125

Ile Asn Arg Ile Tyr Gly Phe Tyr Asp Glu Cys Lys Arg Arg Tyr Asn
 130 135 140

Ile Lys Leu Trp Lys Thr Phe Thr Asp Cys Phe Asn Cys Leu Pro Ile
 145 150 155 160

Ala Ala Ile Ile Asp Glu Lys Ile Phe Cys Met His Gly Gly Leu Ser
 165 170 175

Pro Asp Leu Asn Ser Met Glu Gln Ile Arg Arg Val Met Arg Pro Thr
 180 185 190

Asp Ile Pro Asp Val Gly Leu Leu Cys Asp Leu Leu Trp Ser Asp Pro
 195 200 205

Asp Lys Asp Ile Val Gly Trp Ser Glu Asn Asp Arg Gly Val Ser Phe
 210 215 220

Thr Phe Gly Pro Asp Val Val Asn Arg Phe Leu Gln Lys Gln Asp Met
 225 230 235 240

Glu Leu Ile Cys Arg Ala His Gln Val Val Glu Asp Gly Tyr Glu Phe
 245 250 255

Phe Ser Lys Arg Gln Leu Val Thr Leu Phe Ser Ala Pro Asn Tyr Cys
 260 265 270

Gly Glu Phe Asp Asn Ala Gly Ala Met Met Ser Val Asp Glu Ser Leu
 275 280 285

Leu Cys Ser Phe Gln Ile Leu Lys Pro Ala Gln Lys Ser Leu Pro Arg
 290 295 300

Gln Ala Gly Gly Arg Lys Lys Lys
 305 310

<210> 16
 <211> 939
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 16
 atggactcac aaccagttga cgttgataat atcatcgata gattattgga agtaagagga 60
 tctaaacctg gtcaacaagt tgatctagaa gaaaatgaaa tcagatactt atgttcgaaa 120
 gccagatcta tattcataaa gcaacccatt ttactagagt tagaagcccc aattaaaata 180
 tgtggtgaca ttcattgggca atactatgat ttactacgtc tatttgagta cgggtggattc 240
 ccgccagaat ctaattatct atttttgggt gattatgtcg accgtggtaa acaatcctta 300
 gagactattt gtctattact ggcttacaaa attaagtatc cagaaaactt tttcatttta 360
 agaggggaacc atgaatgtgc ttccattaat agaatttacg ggttttatga tgaatgtaag 420
 agacgttata atatcaaact ttggaaaact ttcacggatt gtttcaattg tttaccaatt 480
 gctgcaatta ttgatgagaa aatcttctgt atgcatggtg gtctctcacc agatttgaat 540
 agtatggaac agatcagaag ggtgatgagg ccaacagata tccccgacgt tggttatta 600
 tgtgacttat tgtggtcaga tccagataaa gatatcgtag gttggagtga aaatgataga 660
 ggtgtttctt tcacttttgg tcttgatgta gtgaacagat ttttacagaa acaagatatg 720
 gagttgattt gcagggccca tcaagttgtg gaagatgggt atgaattctt tagtaaaaga 780
 caattggtga cacttttcag tgctccgaat tattgtggtg aatttgataa cgctgggtgca 840
 atgatgagtg ttgatgaaag tttattatgt tcttttcaaa ttttaaagcc agccccaaaa 900
 agtctaccaa ggcaagctgg gggtagaaag aaaaaataa 939

<210> 17
 <211> 247

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 17

Met Pro Lys Leu Val Leu Val Arg His Gly Gln Ser Glu Trp Asn Glu
 1 5 10 15

Lys Asn Leu Phe Thr Gly Trp Val Asp Val Lys Leu Ser Ala Lys Gly
 20 25 30

Gln Gln Glu Ala Ala Arg Ala Gly Glu Leu Leu Lys Glu Lys Lys Val
 35 40 45

Tyr Pro Asp Val Leu Tyr Thr Ser Lys Leu Ser Arg Ala Ile Gln Thr
 50 55 60

Ala Asn Ile Ala Leu Glu Lys Ala Asp Arg Leu Trp Ile Pro Val Asn
 65 70 75 80

Arg Ser Trp Arg Leu Asn Glu Arg His Tyr Gly Asp Leu Gln Gly Lys
 85 90 95

Asp Lys Ala Glu Thr Leu Lys Lys Phe Gly Glu Glu Lys Phe Asn Thr
 100 105 110

Tyr Arg Arg Ser Phe Asp Val Pro Pro Pro Pro Ile Asp Ala Ser Ser
 115 120 125

Pro Phe Ser Gln Lys Gly Asp Glu Arg Tyr Lys Tyr Val Asp Pro Asn
 130 135 140

Val Leu Pro Glu Thr Glu Ser Leu Ala Leu Val Ile Asp Arg Leu Leu
 145 150 155 160

Pro Tyr Trp Gln Asp Val Ile Ala Lys Asp Leu Leu Ser Gly Lys Thr
 165 170 175

Val Met Ile Ala Ala His Gly Asn Ser Leu Arg Gly Leu Val Lys His
 180 185 190

Leu Glu Gly Ile Ser Asp Ala Asp Ile Ala Lys Leu Asn Ile Pro Thr
 195 200 205

Gly Ile Pro Leu Val Phe Glu Leu Asp Glu Asn Leu Lys Pro Ser Lys
 210 215 220

Pro Ser Tyr Tyr Leu Asp Pro Glu Ala Ala Ala Ala Gly Ala Ala Ala
 225 230 235 240

Val Ala Asn Gln Gly Lys Lys
245

<210> 18
<211> 744
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 18
atgccaaagt tagttttagt tagacacggt caatccgaat ggaacgaaaa gaacttattc 60
accggttggg ttgatgttaa attgtctgcc aagggtcaac aagaagccgc tagagccggt 120
gaattgttga aggaaaagaa ggtctacca gacgtcttgt acacttccaa gttgtccaga 180
gctatccaaa ctgctaacat tgctttggaa aaggctgaca gattatggat tccagtcaac 240
agatcctgga gattgaacga aagacattac ggtgacttac aaggtaagga caaggctgaa 300
actttgaaga agttcgggtga agaaaaattc aacacctaca gaagatcctt cgatgttcca 360
cctcccccaa tcgacgcttc ttctccattc tctcaaaagg gtgatgaaag atacaagtac 420
gttgacccaa atgtcttgcc agaaactgaa tctttggctt tggtcattga cagattggtg 480
ccatactggc aagatgtcat tgccaaggac ttggtgagtg gtaagaccgt catgatcgcc 540
gctcacggta actccttgag aggtttgggt aagcacttgg aaggatatctc tgatgctgac 600
attgctaagt tgaacatccc aactgggtatt ccattggtct tcgaattgga cgaaaacttg 660
aagccatcta agccatctta ctacttggac ccagaagctg ccgctgctgg tgccgctgct 720
gttgccaacc aagtaagaa ataa 744

<210> 19
<211> 327
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 19

Met Ser Ser Ser Glu Asp Glu Asp Asp Lys Phe Leu Tyr Gly Ser Asp
1 5 10 15

Ser Glu Leu Ala Leu Pro Ser Ser Lys Arg Ser Arg Asp Asp Glu Ala
20 25 30

Asp Ala Gly Ala Ser Ser Asn Pro Asp Ile Val Lys Arg Gln Lys Phe
35 40 45

Asp Ser Pro Val Glu Glu Thr Pro Ala Thr Ala Arg Asp Asp Arg Ser
50 55 60

Asp Glu Asp Ile Tyr Ser Asp Ser Ser Asp Asp Ser Asp Ser Asp

65		70		75		80									
Leu	Glu	Val	Ile	Ile	Ser	Leu	Gly	Pro	Asp	Pro	Thr	Arg	Leu	Asp	Ala
			85						90					95	
Lys	Leu	Leu	Asp	Ser	Tyr	Ser	Thr	Ala	Ala	Thr	Ser	Ser	Ser	Lys	Asp
			100					105					110		
Val	Ile	Ser	Val	Ala	Thr	Asp	Val	Ser	Asn	Thr	Ile	Thr	Lys	Thr	Ser
		115					120					125			
Asp	Glu	Arg	Leu	Ile	Thr	Glu	Gly	Glu	Ala	Asn	Gln	Gly	Val	Thr	Ala
	130					135					140				
Thr	Thr	Val	Lys	Ala	Thr	Glu	Ser	Asp	Gly	Asn	Val	Pro	Lys	Ala	Met
145					150					155					160
Thr	Gly	Ser	Ile	Asp	Leu	Asp	Lys	Glu	Gly	Ile	Phe	Asp	Ser	Val	Gly
			165						170					175	
Ile	Thr	Thr	Ile	Asp	Pro	Glu	Val	Leu	Lys	Glu	Lys	Pro	Trp	Arg	Gln
			180					185					190		
Pro	Gly	Ala	Asn	Leu	Ser	Asp	Tyr	Phe	Asn	Tyr	Gly	Phe	Asn	Glu	Phe
		195					200					205			
Thr	Trp	Met	Glu	Tyr	Leu	His	Arg	Gln	Glu	Lys	Leu	Gln	Gln	Asp	Tyr
	210					215					220				
Asn	Pro	Arg	Arg	Ile	Leu	Met	Gly	Leu	Leu	Ser	Leu	Gln	Gln	Gln	Gly
225					230					235					240
Lys	Leu	Asn	Ser	Ala	Asn	Asp	Thr	Asp	Ser	Asn	Leu	Gly	Asn	Ile	Ile
				245					250					255	
Asp	Asn	Asn	Asn	Asn	Val	Asn	Asn	Ala	Asn	Met	Ser	Asn	Leu	Asn	Ser
				260				265					270		
Asn	Met	Gly	Asn	Ser	Met	Ser	Gly	Thr	Pro	Asn	Pro	Pro	Ala	Pro	Pro
		275					280					285			
Met	His	Pro	Ser	Phe	Pro	Pro	Leu	Pro	Met	Phe	Gly	Ser	Phe	Pro	Pro
	290						295				300				
Phe	Pro	Met	Pro	Gly	Met	Met	Pro	Pro	Met	Asn	Gln	Gln	Pro	Asn	Gln
305					310					315					320

Asn Gln Asn Gln Asn Ser Lys
325

<210> 20
<211> 984
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 20
atgagctcca gtgaagacga agacgacaag ttcttgtatg gttccgactc cgaattagca 60
ctaccttcat ctaaacgatc aagagatgat gaagcagacg caggtgcgtc cagtaatcct 120
gatatagtta aaaggcaaaa attcgactct cccgtggaag aaactccagc tactgccaga 180
gatgatcggt ctgatgaaga tatctactct gactcctcag atgacgatag tgattctgac 240
ctagagggtta tcataagtct gggtcctgac cctactaggt tagatgcaaa actactcgat 300
tcttattcta cgcagcgcac atcttcaagc aaagacgtaa ttagcgtagc tacagatgta 360
tccaatacca tcacaaagac atcagatgaa agactaataa cagaaggaga agcaaatcaa 420
gggtgaacgg caacgaccgt aaaagctaca gagagcgatg gaaatgtacc gaaagcaatg 480
actggttcta tagacctgga taaagaggga atctttgata gtggtggcat aacgacaata 540
gatcctgaag tattaagga gaaaccctgg aggcaaccgg gggccaactt aagtgattat 600
ttcaattacg gttttaacga atttacctgg atggagtatt tacatagaca ggaaaaacta 660
caacaagatt ataatcctag gaggatccta atgggcctat tatccctcca acagcaaggg 720
aagttgaatt ccgcgaatga tacagactca aacctcggtat atataattga taacaacaac 780
aacgtaaaca atgcaaatat gtctaactctg aacagtaata tgggtaatag tatgtctgga 840
acaccaaacc ctcccgctcc accaatgcat ccaagcttcc cacccttacc tatgtttggt 900
agctttccac cattccccat gccaggtatg atgccacca tgaaccaaca gcctaatacaa 960
aatcaaaatc aaaattcgaa atga 984

<210> 21
<211> 103
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 21

Met Ser Gly Arg Gly Lys Gly Gly Lys Gly Leu Gly Lys Gly Gly Ala
1 5 10 15

Lys Arg His Arg Lys Ile Leu Arg Asp Asn Ile Gln Gly Ile Thr Lys
20 25 30

Pro Ala Ile Arg Arg Leu Ala Arg Arg Gly Gly Val Lys Arg Ile Ser
35 40 45

Gly Leu Ile Tyr Glu Glu Val Arg Ala Val Leu Lys Ser Phe Leu Glu
50 55 60

Ser Val Ile Arg Asp Ser Val Thr Tyr Thr Glu His Ala Lys Arg Lys
65 70 75 80

Thr Val Thr Ser Leu Asp Val Val Tyr Ala Leu Lys Arg Gln Gly Arg
85 90 95

Thr Leu Tyr Gly Phe Gly Gly
100

<210> 22

<211> 312

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 22

atgtccggtg gaggtaaagg tggtaaagggt ctaggaaaag gtgggtgccaa gcgtcacaga 60

aagattctaa gagataacat tcaaggtatc actaagccag ctatcagaag attagctaga 120

agaggtgggtg tcaagcgtat ttctgggttg atctacgaag aagtcagagc cgtcttgaaa 180

tcctttcttgg aatccgtcat cagggactct gttacttaca ctgaacacgc caagagaaaag 240

actgttactt ctttggatgt tgtttatgct ttgaagagac aaggtagaac cttatatggt 300

ttcgggtggtt aa 312

<210> 23

<211> 132

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 23

Met Ser Gly Gly Lys Gly Gly Lys Ala Gly Ser Ala Ala Lys Ala Ser
1 5 10 15

Gln Ser Arg Ser Ala Lys Ala Gly Leu Thr Phe Pro Val Gly Arg Val
20 25 30

His Arg Leu Leu Arg Arg Gly Asn Tyr Ala Gln Arg Ile Gly Ser Gly
35 40 45

Ala Pro Val Tyr Leu Thr Ala Val Leu Glu Tyr Leu Ala Ala Glu Ile
50 55 60

Leu Glu Leu Ala Gly Asn Ala Ala Arg Asp Asn Lys Lys Thr Arg Ile
65 70 75 80

Gly Ile Gly Met Thr Lys Ala Glu Leu Ile Asn Asn Leu Gly Thr Ile
85 90 95

Ala Lys Ser Gly Thr Lys Ala Phe Met Glu Ala Leu Ser Ala Gly Ala
100 105 110

Asp Val Ser Met Ile Gly Gln Phe Gly Val Gly Phe Tyr Ser Leu Phe
115 120 125

Leu Val Ala Asp Arg Val Gln Val Ile Ser Lys Asn Asn Glu Asp Glu
130 135 140

Gln Tyr Ile Trp Glu Ser Asn Ala Gly Gly Ser Phe Thr Val Thr Leu
145 150 155 160

Asp Glu Val Asn Glu Arg Ile Gly Arg Gly Thr Val Leu Arg Leu Phe
165 170 175

Leu Lys Asp Asp Gln Leu Glu Tyr Leu Glu Glu Lys Arg Ile Lys Glu
180 185 190

Val Ile Lys Arg His Ser Glu Phe Val Ala Tyr Pro Ile Gln Leu Leu
195 200 205

Val Thr Lys Glu Val Glu Lys Glu Val Pro Ile Pro Glu Glu Glu Lys
210 215 220

Lys Asp Glu Glu Lys Lys Asp Glu Asp Asp Lys Lys Pro Lys Leu Glu
225 230 235 240

Glu Val Asp Glu Glu Glu Glu Lys Lys Pro Lys Thr Lys Lys Val
245 250 255

Lys Glu Glu Val Gln Glu Leu Glu Glu Leu Asn Lys Thr Lys Pro Leu
260 265 270

Trp Thr Arg Asn Pro Ser Asp Ile Thr Gln Glu Glu Tyr Asn Ala Phe
275 280 285

Tyr Lys Ser Ile Ser Asn Asp Trp Glu Asp Pro Leu Tyr Val Lys His
290 295 300

Phe Ser Val Glu Gly Gln Leu Glu Phe Arg Ala Ile Leu Phe Ile Pro
305 310 315 320

Lys Arg Ala Pro Phe Asp Leu Phe Glu Ser Lys Lys Lys Lys Asn Asn
325 330 335

Ile Lys Leu Tyr Val Arg Arg Val Phe Ile Thr Asp Glu Ala Glu Asp
 340 345 350

Leu Ile Pro Glu Trp Leu Ser Phe Val Lys Gly Val Val Asp Ser Glu
 355 360 365

Asp Leu Pro Leu Asn Leu Ser Arg Glu Met Leu Gln Gln Asn Lys Ile
 370 375 380

Met Lys Val Ile Arg Lys Asn Ile Val Lys Lys Leu Ile Glu Ala Phe
 385 390 395 400

Asn Glu Ile Ala Glu Asp Ser Glu Gln Phe Asp Lys Phe Tyr Ser Ala
 405 410 415

Phe Ala Lys Asn Ile Lys Leu Gly Val His Glu Asp Thr Gln Asn Arg
 420 425 430

Ala Ala Leu Ala Lys Leu Leu Arg Tyr Asn Ser Thr Lys Ser Val Asp
 435 440 445

Glu Leu Thr Ser Leu Thr Asp Tyr Val Thr Arg Met Pro Glu His Gln
 450 455 460

Lys Asn Ile Tyr Tyr Ile Thr Gly Glu Ser Leu Lys Ala Val Glu Lys
 465 470 475 480

Ser Pro Phe Leu Asp Ala Leu Lys Ala Lys Asn Phe Glu Val Leu Phe
 485 490 495

Leu Thr Asp Pro Ile Asp Glu Tyr Ala Phe Thr Gln Leu Lys Glu Phe
 500 505 510

Glu Gly Lys Thr Leu Val Asp Ile Thr Lys Asp Phe Glu Leu Glu Glu
 515 520 525

Thr Asp Glu Glu Lys Ala Glu Arg Glu Lys Glu Ile Lys Glu Tyr Glu
 530 535 540

Pro Leu Thr Lys Ala Leu Lys Asp Ile Leu Gly Asp Gln Val Glu Lys
 545 550 555 560

Val Val Val Ser Tyr Lys Leu Leu Asp Ala Pro Ala Ala Ile Arg Thr
 565 570 575

Gly Gln Phe Gly Trp Ser Ala Asn Met Glu Arg Ile Met Lys Ala Gln

580	585	590
Ala Leu Arg Asp Ser Ser Met Ser Ser Tyr Met Ser Ser Lys Lys Thr		
595	600	605
Phe Glu Ile Ser Pro Lys Ser Pro Ile Ile Lys Glu Leu Lys Lys Arg		
610	615	620
Val Asp Glu Gly Gly Ala Gln Asp Lys Thr Val Lys Asp Leu Thr Asn		
625	630	635
Leu Leu Phe Glu Thr Ala Leu Leu Thr Ser Gly Phe Ser Leu Glu Glu		
645	650	655
Pro Thr Ser Phe Ala Ser Arg Ile Asn Arg Leu Ile Ser Leu Gly Leu		
660	665	670
Asn Ile Asp Glu Asp Glu Glu Thr Glu Thr Ala Pro Glu Ala Ser Thr		
675	680	685
Glu Ala Pro Val Glu Glu Val Pro Ala Asp Thr Glu Met Glu Glu Val		
690	695	700

Asp
705

<210> 26
 <211> 2000
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 26	
atggctggtg aaacttttga atttcaagct gaaatcactc agttgatgag tttgatcatc	60
aacactgtct attctaaca ggaaattttc ttgagagAAC tgatctctaa cgctccgat	120
gctttagaca aaattagata ccaagctttg tctgatccaa agcaattgga aaccgaacca	180
gatttggttca ttagaatcac cccaaaacca gaagaaaaag ttttggaat cagagattct	240
ggtattggta tgaccaaggc tgaattgatt aacaatttgg gtaccattgc taagtctggt	300
actaaagctt tcatggaagc tctatctgct ggtgccgatg tatccatgat tgggtcaattc	360
ggtgttggtt ttactctttt attcttagtc gccgacagag ttcaagttat ttccaagaac	420
aatgaggacg aacaatatat ttgggaatct aatgccggtg gttctttcac cgttactttg	480
gacgaagtta acgaaagaat tggtagaggt accgtcttga gattattctt gaaagatgac	540
caattggagt acttggaga aaagagaatt aaagaagtca tcaagagaca ttctgaattc	600
gttgettacc ctatccaact tctagtcacc aaggaagtcg aaaaggaagt tccaattcca	660


```

gaagaagaaa agaaagacga ggaaaagaag gatgaagatg acaagaaacc aaaattggaa 720
gaagtcgatg aagaagaaga agaaaagaag ccaaaaacca aaaaagttaa agaagaggtt 780
caagaattag aagagttgaa caagactaag ccattatgga ctagaaaccc atctgatatc 840
actcaagagg aatacaatgc tttctataag tctatttcta acgactggga agaccattg 900
tacgttaagc atttctctgt tgaaggtaa ttggaattta gagctatctt gttcattcca 960
aagagagcac cattcgactt atttgagagt aagaagaaga agaacaatat caagttgtac 1020
gttcgctgtg tcttcacac tgatgaagct gaagacttga ttccagagtg gttatctttc 1080
gtcaagggtg ttgttgactc tgaagattta ccattgaatt tgtccagaga aatgttacia 1140
caaaataaga ttatgaaggc tattagaaag aatattgtca agaaattgat tgaagccttc 1200
aacgaaatcg ctgaagactc cgagcaattt gacaaatttt actctgcctt cgctaagaac 1260
attaagctgg gtgtacatga ggacactcaa aacagagctg ctttagctaa gttgctacgt 1320
tacaattcta ctaaactctgt cgatgaattg acttccttga ctgattacgt tactagaatg 1380
ccagaacacc aaaagaacat ctattacatc accggtgaat ctctaaaggc agtcgaaaag 1440
tctccattct tggacgcctt gaaggctaag aactttgaag ttttgttctt gaccgacca 1500
attgatgaat acgctttcac tcaattgaag gaattcgagg gtaaaacttt gggtgacatt 1560
actaaagatt tcgaattgga agaaacagac gaagaaaaag ctgaaagaga gaaggagatc 1620
aaagaatacg aaccattgac caaggccttg aaggatatct tgggtgacca agtgagagaag 1680
gttggtgttt cttacaaatt gctagatgct ccagctgcca tcagaactgg tcaattcggc 1740
tggctctgta acatggaaag aatcatgaag gctcaagcct tgagagactc tccatgtcc 1800
tcctacatgt cttccaagaa gactttcgaa atttctccaa aatctccaat tattaaggaa 1860
ttgaaaaaga gagttgatga ggggtgtgca caagataaga cgtcaaaga tttgactaac 1920
ttattattcg agaccgcttt gttgacttct ggtttcagtt tggaagaacc aacttctttt 1980
gcatcaagaa taaatagatt 2000

```

<210> 27

<211> 523

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 27

```

Met Ala Ala Ile Arg Asp Tyr Lys Thr Ala Leu Asp Phe Thr Lys Ser
1             5             10             15

```

```

Leu Pro Arg Pro Asp Gly Leu Ser Val Gln Glu Leu Met Asp Ser Lys
20             25             30

```

```

Ile Arg Gly Gly Leu Thr Tyr Asn Asp Phe Leu Ile Leu Pro Gly Leu

```

35		40		45
Val Asp Phe Ala Ser Ser Glu Val Ser Leu Gln Thr Lys Leu Thr Arg				
50		55		60
Asn Ile Thr Leu Asn Ile Pro Leu Val Ser Ser Pro Met Asp Thr Val				
65		70		80
Thr Glu Ser Glu Met Ala Thr Phe Met Ala Leu Leu Gly Gly Ile Gly				
	85		90	95
Phe Ile His His Asn Cys Thr Pro Glu Asp Gln Ala Asp Met Val Arg				
	100		105	110
Arg Val Lys Asn Tyr Glu Asn Gly Phe Ile Asn Asn Pro Ile Val Ile				
	115		120	125
Ser Pro Thr Thr Thr Val Gly Glu Ala Lys Ser Met Lys Glu Lys Tyr				
	130		135	140
Gly Phe Ala Gly Phe Pro Val Thr Thr Asp Gly Lys Arg Asn Ala Lys				
145		150		155
				160
Leu Val Gly Val Ile Thr Ser Arg Asp Ile Gln Phe Val Glu Asp Asn				
	165		170	175
Ser Leu Leu Val Gln Asp Val Met Thr Lys Asn Pro Val Thr Gly Ala				
	180		185	190
Gln Gly Ile Thr Leu Ser Glu Gly Asn Glu Ile Leu Lys Lys Ile Lys				
	195		200	205
Lys Gly Arg Leu Leu Val Val Asp Glu Lys Gly Asn Leu Val Ser Met				
	210		215	220
Leu Ser Arg Thr Asp Leu Met Lys Asn Gln Asn Tyr Pro Leu Ala Ser				
225		230		235
				240
Lys Ser Ala Asn Thr Lys Gln Leu Leu Cys Gly Ala Ser Ile Gly Thr				
	245		250	255
Met Asp Ala Asp Lys Glu Arg Leu Arg Leu Leu Val Lys Ala Gly Leu				
	260		265	270
Asp Val Val Ile Leu Asp Ser Ser Gln Gly Asn Ser Ile Phe Glu Leu				
	275		280	285

Asn Met Leu Lys Trp Val Lys Glu Ser Phe Pro Gly Leu Glu Val Ile
290 295 300

Ala Gly Asn Val Val Thr Arg Glu Gln Ala Ala Asn Leu Ile Ala Ala
305 310 315 320

Gly Ala Asp Gly Leu Arg Ile Gly Met Gly Thr Gly Ser Ile Cys Ile
325 330 335

Thr Gln Glu Val Met Ala Cys Gly Arg Pro Gln Gly Thr Ala Val Tyr
340 345 350

Asn Val Cys Glu Phe Ala Asn Gln Phe Gly Val Pro Cys Met Ala Asp
355 360 365

Gly Gly Val Gln Asn Ile Gly His Ile Thr Lys Ala Leu Ala Leu Gly
370 375 380

Ser Ser Thr Val Met Met Gly Gly Met Leu Ala Gly Thr Thr Glu Ser
385 390 395 400

Pro Gly Glu Tyr Phe Tyr Gln Asp Gly Lys Arg Leu Lys Ala Tyr Arg
405 410 415

Gly Met Gly Ser Ile Asp Ala Met Gln Lys Thr Gly Thr Lys Gly Asn
420 425 430

Ala Ser Thr Ser Arg Tyr Phe Ser Glu Ser Asp Ser Val Leu Val Ala
435 440 445

Gln Gly Val Ser Gly Ala Val Val Asp Lys Gly Ser Ile Lys Lys Phe
450 455 460

Ile Pro Tyr Leu Tyr Asn Gly Leu Gln His Ser Cys Gln Asp Ile Gly
465 470 475 480

Cys Arg Ser Leu Thr Leu Leu Lys Asn Asn Val Gln Arg Gly Lys Val
485 490 495

Arg Phe Glu Phe Arg Thr Ala Ser Ala Gln Leu Glu Gly Gly Val His
500 505 510

Asn Leu His Ser Tyr Glu Lys Arg Leu His Asn
515 520

<210> 28
<211> 1572
<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 28

atggccgcca ttagagacta caagaccgca ctagacttta ccaagagcct accaagaccg 60
gatggtttgt cagtgcagga attgatggac tccaagatca gaggtgggtt gacttataac 120
gatttttttaa tcttaccagg tttagtcgat tttgcgtcct ctgaagttag cctacagacc 180
aagctaacca ggaatattac tttaaattatt ccattagttt cctctccaat ggacacgggtg 240
acagagtcag aaatggccac ttttatggct ctggtgggtg gtatcggttt cattcaccat 300
aactgtaccc cagaggacca agctgacatg gtcagaagag tcaagaacta tgaaaatggg 360
tttattaaca accctatagt gattttctcca actacgaccg ttggtgaagc taagagcatg 420
aaggaaaagt atggatttgc aggtttccct gtcacgacag atggaaagag aaatgcaaag 480
ttggtgggtg tcatcacctc tcgtgatata caattcgttg aggacaactc tttactcgtt 540
caggatgtca tgaccaaaaa ccctgttacc ggcgcaaacg gtatcacatt atcagaagggt 600
aacgaaattc taaagaaaat caaaaagggt aggtatttgg ttggtgatga aaagggtaac 660
ttagtttcta tgctttcccg aactgattta atgaaaaatc agaactaccc attagcgtcc 720
aaatctgcca acaccaagca actgttatgt ggtgcttcta ttgggactat ggacgctgat 780
aaagaaagac taagattatt ggtaaaagct ggggttggatg tcgtcatatt ggattcatcc 840
caaggaaact ctatcttcga attgaacatg ctcaagtggg tcaaagagag tttcccagggt 900
ctggaagtca tcgctggtaa cgttgtcacc agggaacaag ctgccaattt gattgctgcc 960
ggtgcggacg gtttgagaat tggatatggga actggctcta tttgtattac ccaagaagtt 1020
atggcttgtg gtaggccaca aggtacagcc gtctacaatg tgtgtgaatt tgctaaccaa 1080
ttcgggtgtc catgtatggc tgatgggtgg gttcaaaaca ttggtcatat taccaaagct 1140
ttggctcttg gttcttctac tgttatgatg ggtgggtatg tggccggtac taccgaatca 1200
ccagggtgaat atttctatca agatggtaaa agattgaagg cgtatcgtgg tatgggctcc 1260
attgacgcca tgcaaaagac tggtaacaaa ggtaatgcat ctacctcccg ttacttttcc 1320
gaatcagaca gtgttttggg cgcacaagggt gtctccggtg ctgtcgttga caaaggatcc 1380
attaagaaat ttattccata tttgtacaat ggtttacaac attcctgtca agacatcggc 1440
tgtaggtcgc taactttatt aaagaataat gttcaaaggg gtaaagttag atttgaattc 1500
agaaccgctt ctgctcaact agaaggtggc gttcataatt tacattctta cgaaaagcgt 1560
ttacataact ga 1572

<210> 29

<211> 524

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 29

Met Ser Ala Ala Pro Leu Asp Tyr Lys Lys Ala Leu Glu His Leu Lys
 1 5 10 15

Thr Tyr Ser Ser Lys Asp Gly Leu Ser Val Gln Glu Leu Met Asp Ser
 20 25 30

Thr Thr Arg Gly Gly Leu Thr Tyr Asn Asp Phe Leu Val Leu Pro Gly
 35 40 45

Leu Val Asn Phe Pro Ser Ser Ala Val Ser Leu Gln Thr Lys Leu Thr
 50 55 60

Lys Lys Ile Thr Leu Asn Thr Pro Phe Val Ser Ser Pro Met Asp Thr
 65 70 75 80

Val Thr Glu Ala Asp Met Ala Ile Tyr Met Ala Leu Leu Gly Gly Ile
 85 90 95

Gly Phe Ile His His Asn Cys Thr Pro Lys Glu Gln Ala Ser Met Val
 100 105 110

Lys Lys Val Lys Met Phe Glu Asn Gly Phe Ile Asn Ser Pro Ile Val
 115 120 125

Ile Ser Pro Thr Thr Thr Val Gly Glu Val Lys Val Met Lys Arg Lys
 130 135 140

Phe Gly Phe Ser Gly Phe Pro Val Thr Glu Asp Gly Lys Cys Pro Gly
 145 150 155 160

Lys Leu Val Gly Leu Val Thr Ser Arg Asp Ile Gln Phe Leu Glu Asp
 165 170 175

Asp Ser Leu Val Val Ser Glu Val Met Thr Lys Asn Pro Val Thr Gly
 180 185 190

Ile Lys Gly Ile Thr Leu Lys Glu Gly Asn Glu Ile Leu Lys Gln Thr
 195 200 205

Lys Lys Gly Lys Leu Leu Ile Val Asp Asp Asn Gly Asn Leu Val Ser
 210 215 220

Met Leu Ser Arg Ala Asp Leu Met Lys Asn Gln Asn Tyr Pro Leu Ala
 225 230 235 240

Ser Lys Ser Ala Thr Thr Lys Gln Leu Leu Cys Gly Ala Ala Ile Gly

245	250	255
Thr Ile Glu Ala Asp Lys Glu Arg Leu Arg Leu Leu Val Glu Ala Gly		
260	265	270
Leu Asp Val Val Ile Leu Asp Ser Ser Gln Gly Asn Ser Val Phe Gln		
275	280	285
Leu Asn Met Ile Lys Trp Ile Lys Glu Thr Phe Pro Asp Leu Glu Ile		
290	295	300
Ile Ala Gly Asn Val Ala Thr Arg Glu Gln Ala Ala Asn Leu Ile Ala		
305	310	315
Ala Gly Ala Asp Gly Leu Arg Ile Gly Met Gly Ser Gly Ser Ile Cys		
325	330	335
Ile Thr Gln Glu Val Met Ala Cys Gly Arg Pro Gln Gly Thr Ala Val		
340	345	350
Tyr Asn Val Cys Gln Phe Ala Asn Gln Phe Gly Val Pro Cys Met Ala		
355	360	365
Asp Gly Gly Val Gln Asn Ile Gly His Ile Thr Lys Ala Leu Ala Leu		
370	375	380
Gly Ser Ser Thr Val Met Met Gly Gly Met Leu Ala Gly Thr Thr Glu		
385	390	395
Ser Pro Gly Glu Tyr Phe Tyr Lys Asp Gly Lys Arg Leu Lys Ala Tyr		
405	410	415
Arg Gly Met Gly Ser Ile Asp Ala Met Gln Lys Thr Gly Asn Lys Gly		
420	425	430
Asn Ala Ser Thr Ser Arg Tyr Phe Ser Glu Ser Asp Ser Val Leu Val		
435	440	445
Ala Gln Gly Val Ser Gly Ala Val Val Asp Lys Gly Ser Ile Lys Lys		
450	455	460
Phe Ile Pro Tyr Leu Tyr Asn Gly Leu Gln His Ser Cys Gln Asp Ile		
465	470	475
Gly Cys Glu Ser Leu Thr Ser Leu Lys Glu Asn Val Gln Asn Gly Glu		
485	490	495

Val Arg Phe Glu Phe Arg Thr Ala Ser Ala Gln Leu Glu Gly Gly Val
 500 505 510

His Asn Leu His Ser Tyr Glu Lys Arg Leu Tyr Asn
 515 520

<210> 30

<211> 1575

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 30

atgagtgctg ctccattgga ttacaaaaag gctttagaac atttgaagac atacagttcc 60
 aaggacggtc tttctgtcca ggaattgatg gactccacaa ctagagggtg gttgacctac 120
 aatgattttt tggctctacc aggcttggtc aatttcccat cttctgctgt cagtttgcaa 180
 accaaattga ccaagaaaat cactttgaac actccttttg tctcttctcc tatggacaca 240
 gtcactgaag ctgatatggc tatttatatg gctttattgg gtggtattgg tttcatccat 300
 cacaactgta ctccaaagga acaggcttcc atgggtcaaga aagttaaaat gtttgaaaac 360
 ggtttcatca attctccaat agtaatttct ccaaccacca ctggttggtga agttaagggt 420
 atgaagagaa agtttggttt ctctggcttc ccagttactg aagacggtaa gtgtccagggt 480
 aagttagttg ggtagtcac ctctcgtgat atacaattct tagaagatga ttctttgggt 540
 gttttgaaag tcatgactaa aaatccagtt actggcatca agggatttac tttgaaagaa 600
 ggtaatgaaa tcctaaaaca aaccaagaaa ggtaaattgc ttatcgttga tgataacgggt 660
 aacctcgttt ccatgttgctc aagagcggat ttgatgaaga atcaaaacta cccgttagct 720
 tctaaatccg ccaccaccaa gcaattgcta tgtggtgctg caattggtac tatcgaagct 780
 gataaggaaa gattaagact attagtcgaa gcaggtttgg atggtgttat cttagattcc 840
 tctcaaggta actctgtttt ccaattgaac atgatcaaat ggattaaaga aactttccca 900
 gatttggaag tcattgctgg taacgttgcc accagagaac aagctgctaa cttgattgct 960
 gccggtgccg atggtttaag aattggtatg ggttctgggt ctatttgtat cactcaagaa 1020
 gttatggctt gtggtagacc acaaggtaca gctgtctaca acgtttgtca atttgctaac 1080
 caatttggcg ttccatgtat ggctgatggg ggtgtccaaa acattggcca catcaccaag 1140
 gctttggccc ttggttcctc caccgtcatg atgggtggtg tgtagctgg tactaccgag 1200
 tcaccagggtg aatacttcta caaggatggg aagagattga aggcttatcg tggtagggg 1260
 tccattgacg ccatgcaaaa gaccggtaac aagggtaatg cctctacttc tcgttatttc 1320
 tctgagtcag acagtgtctt gggtgcacaa ggtgtttctg gggctgtcgt agacaaagggt 1380
 tccatcaaga agtttattcc atattgtac aacggtttac aacattcttg tcaagacatt 1440
 gggtgcgaat ctctaacttc attgaaagag aatgttcaaa atgggtgaagt tagatttgaa 1500

ttcagaaccg cttctgctca attagaaggt ggtgtccata atttgactc ctatgaaaaa 1560
 cgtctatatac attga 1575

<210> 31
 <211> 767
 <212> PRT
 <213> *Saccharomyces cerevisiae*
 <400> 31

Met Val Gln Ser Ala Val Leu Gly Phe Pro Arg Ile Gly Pro Asn Arg
 1 5 10 15

Glu Leu Lys Lys Ala Thr Glu Gly Tyr Trp Asn Gly Lys Ile Thr Val
 20 25 30

Asp Glu Leu Phe Lys Val Gly Lys Asp Leu Arg Thr Gln Asn Trp Lys
 35 40 45

Leu Gln Lys Glu Ala Gly Val Asp Ile Ile Pro Ser Asn Asp Phe Ser
 50 55 60

Phe Tyr Asp Gln Val Leu Asp Leu Ser Leu Leu Phe Asn Val Ile Pro
 65 70 75 80

Asp Arg Tyr Thr Lys Tyr Asp Leu Ser Pro Ile Asp Thr Leu Phe Ala
 85 90 95

Met Gly Arg Gly Leu Gln Arg Lys Ala Thr Glu Thr Glu Lys Ala Val
 100 105 110

Asp Val Thr Ala Leu Glu Met Val Lys Trp Phe Asp Ser Asn Tyr His
 115 120 125

Tyr Val Arg Pro Thr Phe Ser Lys Thr Thr Gln Phe Lys Leu Asn Gly
 130 135 140

Gln Lys Pro Val Asp Glu Phe Leu Glu Ala Lys Glu Leu Gly Ile His
 145 150 155 160

Thr Arg Pro Val Leu Leu Gly Pro Val Ser Tyr Leu Phe Leu Gly Lys
 165 170 175

Ala Asp Lys Asp Ser Leu Asp Leu Glu Pro Leu Ser Leu Leu Glu Gln
 180 185 190

Leu Leu Pro Leu Tyr Thr Glu Ile Leu Ser Lys Leu Ala Ser Ala Gly
 195 200 205

Ala Thr Glu Val Gln Ile Asp Glu Pro Val Leu Val Leu Asp Leu Pro
 210 215 220

Ala Asn Ala Gln Ala Ala Ile Lys Lys Ala Tyr Thr Tyr Phe Gly Glu
 225 230 235 240

Gln Ser Asn Leu Pro Lys Ile Thr Leu Ala Thr Tyr Phe Gly Thr Val
 245 250 255

Val Pro Asn Leu Asp Ala Ile Lys Gly Leu Pro Val Ala Ala Leu His
 260 265 270

Val Asp Phe Val Arg Ala Pro Glu Gln Phe Asp Glu Val Val Ala Ala
 275 280 285

Ile Gly Asn Lys Gln Thr Leu Ser Val Gly Ile Val Asp Gly Arg Asn
 290 295 300

Ile Trp Lys Asn Asp Phe Lys Lys Ser Ser Ala Ile Val Asn Lys Ala
 305 310 315 320

Ile Glu Lys Leu Gly Ala Asp Arg Val Val Val Ala Thr Ser Ser Ser
 325 330 335

Leu Leu His Thr Pro Val Asp Leu Asn Asn Glu Thr Lys Leu Asp Ala
 340 345 350

Glu Ile Lys Gly Phe Phe Ser Phe Ala Thr Gln Lys Leu Asp Glu Val
 355 360 365

Val Val Ile Thr Lys Asn Val Ser Gly Gln Asp Val Ala Ala Ala Leu
 370 375 380

Glu Ala Asn Ala Lys Ser Val Glu Ser Arg Gly Lys Ser Lys Phe Ile
 385 390 395 400

His Asp Ala Ala Val Lys Ala Arg Val Ala Ser Ile Asp Glu Lys Met
 405 410 415

Ser Thr Arg Ala Ala Pro Phe Glu Gln Arg Leu Pro Glu Gln Gln Lys
 420 425 430

Val Phe Asn Leu Pro Leu Phe Pro Thr Thr Thr Ile Gly Ser Phe Pro
 435 440 445

Gln Thr Lys Asp Ile Arg Ile Asn Arg Asn Lys Phe Asn Lys Gly Thr

450		455		460
Ile Ser Ala Glu Glu Tyr Glu Lys Phe Ile Asn Ser Glu Ile Glu Lys				
465		470	475	480
Val Ile Arg Phe Gln Glu Glu Ile Gly Leu Asp Val Leu Val His Gly				
	485		490	495
Glu Pro Glu Arg Asn Asp Met Val Gln Tyr Phe Gly Glu Gln Ile Asn				
	500		505	510
Gly Tyr Ala Phe Thr Val Asn Gly Trp Val Gln Ser Tyr Gly Ser Arg				
	515		520	525
Tyr Val Arg Pro Pro Ile Ile Val Gly Asp Leu Ser Arg Pro Lys Ala				
	530		535	540
Met Ser Val Lys Glu Ser Val Tyr Ala Gln Ser Ile Thr Ser Lys Pro				
	545		550	555
Val Lys Gly Met Leu Thr Gly Pro Ile Thr Cys Leu Arg Trp Ser Phe				
	565		570	575
Pro Arg Asp Asp Val Asp Gln Lys Thr Gln Ala Met Gln Leu Ala Leu				
	580		585	590
Ala Leu Arg Asp Glu Val Asn Asp Leu Glu Ala Ala Gly Ile Lys Val				
	595		600	605
Ile Gln Val Asp Glu Pro Ala Leu Arg Glu Gly Leu Pro Leu Arg Glu				
	610		615	620
Gly Thr Glu Arg Ser Ala Tyr Tyr Thr Trp Ala Ala Glu Ala Phe Arg				
	625		630	635
Val Ala Thr Ser Gly Val Ala Asn Lys Thr Gln Ile His Ser His Phe				
	645		650	655
Cys Tyr Ser Asp Leu Asp Pro Asn His Ile Lys Ala Leu Asp Ala Asp				
	660		665	670
Val Val Ser Ile Glu Phe Ser Lys Lys Asp Asp Ala Asn Tyr Ile Ala				
	675		680	685
Glu Phe Lys Asn Tyr Pro Asn His Ile Gly Leu Gly Leu Phe Asp Ile				
	690		695	700

His Ser Pro Arg Ile Pro Ser Lys Asp Glu Phe Ile Ala Lys Ile Ser
705 710 715 720

Thr Ile Leu Lys Ser Tyr Pro Ala Glu Lys Phe Trp Val Asn Pro Asp
725 730 735

Cys Gly Leu Lys Thr Arg Gly Trp Glu Glu Thr Arg Leu Ser Leu Thr
740 745 750

His Met Val Glu Ala Ala Lys Tyr Phe Arg Glu Gln Tyr Lys Asn
755 760 765

<210> 32

<211> 2304

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 32

atgggttcaat ctgctgtctt agggttccca agaatcggtc caaacagaga attaaagaag	60
gccactgaag gttactggaa cggtaaaatc actgtcgatg aattattcaa ggtcggtaag	120
gatttgagaa ctcaaaactg gaagttgcaa aaggaggctg gtgttgatat catcccatcc	180
aatgacttct ccttttacga ccaagttttg gatttgtctt tgttgttcaa tgtcattcca	240
gaccgttaca ctaagtacga tctatctcca atcgacactt tgtttgctat gggtagagggt	300
ttacaaagaa aggccactga aactgaaaag gctgtcgacg tctactgctt ggaaatgggt	360
aaatgggtcg actctaacta ccattacgtt agaccaactt tctccaagac cactcaattt	420
aagttgaacg gccaaaagcc agttgacgaa tttttggaag ccaaggagtt aggtattcat	480
actagacctg tcttgttggg tccagtttct tacttattct tgggtaaggc tgacaaggat	540
tctctagatt tggaaccatt gtccctattg gaacaattgt tgcctctata cactgaaatc	600
ctatctaaat tggtctctgc tggtgccact gaagttcaaa ttgacgaacc tgtcttagtt	660
ttggacttgc ctgccaacgc ccaagccgcc attaagaagg cttacactta cttcggtgaa	720
caaagcaatc taccaaagat tactttggct acttacttcg gtaccgttgt ccctaactta	780
gacgccatca agggcttgcc agttgctgcc ttacacgttg actttgttag agctccagaa	840
caatttgatg aagtcgttgc cgccattggg aacaaacaaa ccttgtccgt tggatttggt	900
gatggtagaa acatttgga gaattgattc aagaagtctt ccgctatcgt taacaaggct	960
attgaaaagt tgggtgctga cagagtcgtt gttgccactt cttcttctct attgcacaca	1020
ccagttgatt tgaacaacga aaccaagttg gacgctgaaa tcaagggtt tttctctttc	1080
gccactcaaa aattggatga agttgttgtg atcaccaaga acgtttccgg tcaagacgtt	1140
gctgctgccc tagaagctaa cgctaaatct gttgaatcca gaggtaaatc caagtttate	1200
cacgatgctg ccgttaaggc cagagttgcc tctatcgacg aaaaaatgtc tactagagca	1260

```

gctccatttg aacaaagatt gcctgaacaa caaaaagtct tcaacttgcc attgttccca 1320
acaacaacta ttgggttcctt ccctcaaacc aaggacatca gaattaacag aaacaaattc 1380
aacaagggca ccatctctgc tgaagaatat gaaaaattca tcaattctga aattgaaaag 1440
gtcatcagat tccaagaaga aattggtttg gatgtcttag tccacggtga accagaaaga 1500
aacgatatgg ttcaatactt cgggtgaacaa atcaacgggt atgctttcac tgtaacgggt 1560
tgggttcaat cttacgggtc cagatatgtc agaccaccaa ttattgttg tgacttgtcc 1620
agaccaaagg ctatgtccgt caaggaatct gtttacgctc aatccatcac ttctaagcca 1680
gtaaagggta tgttgactgg tccaattacc tgtttgagat ggtctttccc aagagacgat 1740
gtcgacaaaa aaactcaagc tatgcaatta gctttggctt tgagagatga agtcaatgat 1800
ttggaagctg ccggtatcaa ggttatccaa gttgatgaac cagctttaag agaaggttta 1860
ccattgagag aaggctactga gagatctgct tactacacct gggctgccga agctttcaga 1920
gttgctactt ctggtgttgc taacaagact caaatacact ctcatctctg ttactctgac 1980
ttggatccaa accatatcaa ggctttggat gctgatgttg tttccatcga attctctaag 2040
aaggacgatg ctaactacat tgetgaattc aaaaactatc caaaccacat tggcttaggt 2100
ttattcgata ttcattctcc aagaattcca tcaaaggatg aatttatcgc caagatttca 2160
accatcttga agagctaccc agctgaaaag ttctgggtta acccagattg tggtttgaag 2220
actagaggct gggaagaaac tagattgtct ttgactcata tggtcgaagc cgccaagtac 2280
ttccgtgaac aatacaagaa ttaa 2304

```

```

<210> 33
<211> 577
<212> PRT
<213> Saccharomyces cerevisiae

```

```

<400> 33

```

```

Met Ala Asp Ile Thr Asp Lys Thr Ala Glu Gln Leu Glu Asn Leu Asn
1           5           10          15

```

```

Ile Gln Asp Asp Gln Lys Gln Ala Ala Thr Gly Ser Glu Ser Gln Ser
20          25          30

```

```

Val Glu Asn Ser Ser Ala Ser Leu Tyr Val Gly Asp Leu Glu Pro Ser
35          40          45

```

```

Val Ser Glu Ala His Leu Tyr Asp Ile Phe Ser Pro Ile Gly Ser Val
50          55          60

```

```

Ser Ser Ile Arg Val Cys Arg Asp Ala Ile Thr Lys Thr Ser Leu Gly
65          70          75          80

```

Tyr Ala Tyr Val Asn Phe Asn Asp His Glu Ala Gly Arg Lys Ala Ile
 85 90 95

Glu Gln Leu Asn Tyr Thr Pro Ile Lys Gly Arg Leu Cys Arg Ile Met
 100 105 110

Trp Ser Gln Arg Asp Pro Ser Leu Arg Lys Lys Gly Ser Gly Asn Ile
 115 120 125

Phe Ile Lys Asn Leu His Pro Asp Ile Asp Asn Lys Ala Leu Tyr Asp
 130 135 140

Thr Phe Ser Val Phe Gly Asp Ile Leu Ser Ser Lys Ile Ala Thr Asp
 145 150 155 160

Glu Asn Gly Lys Ser Lys Gly Phe Gly Phe Val His Phe Glu Glu Glu
 165 170 175

Gly Ala Ala Lys Glu Ala Ile Asp Ala Leu Asn Gly Met Leu Leu Asn
 180 185 190

Gly Gln Glu Ile Tyr Val Ala Pro His Leu Ser Arg Lys Glu Arg Asp
 195 200 205

Ser Gln Leu Glu Glu Thr Lys Ala His Tyr Thr Asn Leu Tyr Val Lys
 210 215 220

Asn Ile Asn Ser Glu Thr Thr Asp Glu Gln Phe Gln Glu Leu Phe Ala
 225 230 235 240

Lys Phe Gly Pro Ile Val Ser Ala Ser Leu Glu Lys Asp Ala Asp Gly
 245 250 255

Lys Leu Lys Gly Phe Gly Phe Val Asn Tyr Glu Lys His Glu Asp Ala
 260 265 270

Val Lys Ala Val Glu Ala Leu Asn Asp Ser Glu Leu Asn Gly Glu Lys
 275 280 285

Leu Tyr Val Gly Arg Ala Gln Lys Lys Asn Glu Arg Met His Val Leu
 290 295 300

Lys Lys Gln Tyr Glu Ala Tyr Arg Leu Glu Lys Met Ala Lys Tyr Gln
 305 310 315 320

Gly Val Asn Leu Phe Val Lys Asn Leu Asp Asp Ser Val Asp Asp Glu

Ala Ala Tyr Glu Ser Phe Lys Lys Glu Gln Glu Gln Gln Thr Glu Gln
565 570 575

Ala

<210> 34
<211> 1734
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 34
atggctgata ttactgataa gacagctgaa caattggaaa acttgaatat tcaagatgac 60
caaaagcaag ccgccactgg ttcagaaagc caatctgttg aaaactcttc tgcattcatta 120
tatgttggtg acttagaacc ttctgtttcc gaagcccact tatatgatat cttctctcca 180
atcggttcag tctcctccat tcgtgtctgt cgtgatgcca tactaagac ctctttgggc 240
tatgcttatg ttaactttaa cgaccatgaa gccggcagaa aagcaattga gcaattgaac 300
tacactccaa tcaagggtag attatgccgt attatgtggc ctcaacgtga cccatcattg 360
agaaagaagg gttctggtaa catctttatc aagaacttgc accctgatat tgacaacaag 420
gctttgtatg acactttctc tgtgtttggc gacatcttgc ccagcaagat tgccaccgac 480
gaaaacggaa aatccaaggg ttttgggttt gttcacttcg aagaagaagg tgctgccaag 540
gaagctattg atgctttgaa tggatgctg ttgaacggc aagaaattta tgttgctcct 600
cacttgctca gaaaggaacg tgactctcaa ttggaagaga ctaaggcaca ttactaac 660
ctttatgtga aaaacatcaa ctccgaaact actgacgaac aattccaaga attgtttgcc 720
aaatttggtc caattgtttc tgcctctttg gaaaaggatg ctgatggaaa attgaagggt 780
ttcgggtttg ttaactacga aaagcatgaa gacgctgtga aagctgttga agctttgaat 840
gactctgaac taaatggaga aaagtatac gttggctcgtg cccaaaagaa gaatgaacgt 900
atgcatgtct tgaagaagca atacgaagct tacagattgg aaaaaatggc caagtaccaa 960
ggtgttaatt tgtttgttaa gaacttagat gacagcgttg atgacgaaaa gttggaagaa 1020
gaatttgctc catatggtac tatcacttct gcaaagggtta tgagaaccga aaacggtaag 1080
tctaagggtt ttggttttgt ttgtttctca actccagagg aagctactaa ggccattaca 1140
gaaaagaacc aacaaattgt tgctggtaag ccattatacg ttgccattgc tcaaagaaaa 1200
gacgtaagac gttctcaatt ggctcaacaa atccaagcca gaaatcaaat gagataccag 1260
caagctactg ctgccgctgc cgcgcgcgt gccggtatgc cagggtcaatt catgcctcca 1320
atgttctatg gtgttatgcc accaagaggt gttccattca acggtccaaa cccacaacaa 1380
atgaacccaa tgggcggtat gccaaagaac ggcattgccac ctcaatttag aaatgggtccg 1440
gtttacggcg tccccccaca aggtggtttc ccaagaaatg ccaacgataa caaccaattt 1500
tatcaacaaa agcaaagaca agctttgggt gaacaattat acaagaaggt ttctgctaag 1560
acttcaaatg aagaagcagc tggtaaaatt actggtatga ttttggattt gccacctcaa 1620

gaggtcttcc cattgttga aagtgatgaa ttgttcgaac aacactacaa agaagcttct 1680
 gctgcctatg agtctttcaa aaaggagcaa gaacaacaaa ctgagcaagc ttaa 1734

<210> 35
 <211> 568
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 35

Met Ser Ser Gln Lys Val Phe Gly Ile Thr Gly Pro Val Ser Thr Val
 1 5 10 15

Gly Ala Thr Ala Ala Glu Asn Lys Leu Asn Asp Ser Leu Ile Gln Glu
 20 25 30

Leu Lys Lys Glu Gly Ser Phe Glu Thr Glu Gln Glu Thr Ala Asn Arg
 35 40 45

Val Gln Val Leu Lys Ile Leu Gln Glu Leu Ala Gln Arg Phe Val Tyr
 50 55 60

Glu Val Ser Lys Lys Lys Asn Met Ser Asp Gly Met Ala Arg Asp Ala
 65 70 75 80

Gly Gly Lys Ile Phe Thr Tyr Gly Ser Tyr Arg Leu Gly Val His Gly
 85 90 95

Pro Gly Ser Asp Ile Asp Thr Leu Val Val Val Pro Lys His Val Thr
 100 105 110

Arg Glu Asp Phe Phe Thr Val Phe Asp Ser Leu Leu Arg Glu Arg Lys
 115 120 125

Glu Leu Asp Glu Ile Ala Pro Val Pro Asp Ala Phe Val Pro Ile Ile
 130 135 140

Lys Ile Lys Phe Ser Gly Ile Ser Ile Asp Leu Ile Cys Ala Arg Leu
 145 150 155 160

Asp Gln Pro Gln Val Pro Leu Ser Leu Thr Leu Ser Asp Lys Asn Leu
 165 170 175

Leu Arg Asn Leu Asp Glu Lys Asp Leu Arg Ala Leu Asn Gly Thr Arg
 180 185 190

Val Thr Asp Glu Ile Leu Glu Leu Val Pro Lys Pro Asn Val Phe Arg
 195 200 205

Ile Ala Leu Arg Ala Ile Lys Leu Trp Ala Gln Arg Arg Ala Val Tyr
 210 215 220

Ala Asn Ile Phe Gly Phe Pro Gly Gly Val Ala Trp Ala Met Leu Val
 225 230 235 240

Ala Arg Ile Cys Gln Leu Tyr Pro Asn Ala Cys Ser Ala Val Ile Leu
 245 250 255

Asn Arg Phe Phe Ile Ile Leu Ser Glu Trp Asn Trp Pro Gln Pro Val
 260 265 270

Ile Leu Lys Pro Ile Glu Asp Gly Pro Leu Gln Val Arg Val Trp Asn
 275 280 285

Pro Lys Ile Tyr Ala Gln Asp Arg Ser His Arg Met Pro Val Ile Thr
 290 295 300

Pro Ala Tyr Pro Ser Met Cys Ala Thr His Asn Ile Thr Glu Ser Thr
 305 310 315 320

Lys Lys Val Ile Leu Gln Glu Phe Val Arg Gly Val Gln Ile Thr Asn
 325 330 335

Asp Ile Phe Ser Asn Lys Lys Ser Trp Ala Asn Leu Phe Glu Lys Asn
 340 345 350

Asp Phe Phe Phe Arg Tyr Lys Phe Tyr Leu Glu Ile Thr Ala Tyr Thr
 355 360 365

Arg Gly Ser Asp Glu Gln His Leu Lys Trp Ser Gly Leu Val Glu Ser
 370 375 380

Lys Val Arg Leu Leu Val Met Lys Leu Glu Val Leu Ala Gly Ile Lys
 385 390 395 400

Ile Ala His Pro Phe Thr Lys Pro Phe Glu Ser Ser Tyr Cys Cys Pro
 405 410 415

Thr Glu Asp Asp Tyr Glu Met Ile Gln Asp Lys Tyr Gly Ser His Lys
 420 425 430

Thr Glu Thr Ala Leu Asn Ala Leu Lys Leu Val Thr Asp Glu Asn Lys
 435 440 445

Glu Glu Glu Ser Ile Lys Asp Ala Pro Lys Ala Tyr Leu Ser Thr Met

450 455 460
 Tyr Ile Gly Leu Asp Phe Asn Ile Glu Asn Lys Lys Glu Lys Val Asp
 465 470 475 480

 Ile His Ile Pro Cys Thr Glu Phe Val Asn Leu Cys Arg Ser Phe Asn
 485 490 495

 Glu Asp Tyr Gly Asp His Lys Val Phe Asn Leu Ala Leu Arg Phe Val
 500 505 510

 Lys Gly Tyr Asp Leu Pro Asp Glu Val Phe Asp Glu Asn Glu Lys Arg
 515 520 525

 Pro Ser Lys Lys Ser Lys Arg Lys Asn Leu Asp Ala Arg His Glu Thr
 530 535 540

 Val Lys Arg Ser Lys Ser Asp Ala Ala Ser Gly Asp Asn Ile Asn Gly
 545 550 555 560

 Thr Thr Ala Ala Val Asp Val Asn
 565

<210> 36

<211> 1707

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 36

atgagctctc aaaagggtttt tgggtattact ggacctgttt ccacctgtgg cgccacagca 60
 gcagaaaata aattaaatga tagtttaatc caagaactga aaaaggaagg atcgttcgaa 120
 acagagcaag aaactgccaa tagggtacaa gtgttgaaaa tattgcagga attggcacia 180
 agatttggtt atgaagtatc gaagaagaaa aatatgtcag acgggatggc aagggatgct 240
 ggtgggaaga tttttacgta tgggtcctat agactaggag tccatgggcc tggtagtgat 300
 atcgatactt tggtagttgt tccaaaacat gtaactcggg aagatTTTTT tacggtatTT 360
 gattcactac tgagagagag gaaggaactg gatgaaattg cacctgtacc tgatgcgttt 420
 gtcccgatta tcaagataaa gttcagtggg atttctatcg atttaactctg tgcaactcta 480
 gaccaacctc aagtgccttt atccttgact ttatcagata aaaatctact gcgaaatcta 540
 gacgagaagg acttgagagc tttgaatggg accagagtaa cagatgagat attagaactg 600
 gtaccaaagc cgaatgtttt cagaatcgct ttaagagcta ttaagctatg ggcccaaaga 660
 agggctgttt atgctaatat ttttggtttt cctgggtggg tggcttgggc catgctagtg 720
 gctagaattt gtcaactata ccctaacgcc tgtagcgcag ttatattgaa cagatTTTTc 780

```

atcattttgt cggaatggaa ttggccacaa cctgttatct tgaaaccaat tgaggatggc      840
ccgttacaag ttcgtgtatg gaatccaaag atatatgcc aagacaggtc tcatagaatg      900
cccgtcatta caccagctta tccatcaatg tgtgctaccc ataacatcac ggaatctact      960
aaaaaagtca ttttacagga attcgtaaga ggcgttcaaa ttacgaatga tattttttcc     1020
aataagaagt cctggggccaa tttattcgaa aaaaacgatt ttttctttcg atacaagttc     1080
tatttagaaa ttactgcata tacaaggggc agtgacgagc agcattttaa atggagtggt     1140
cttggtgaaa gtaaggtaag gcttctagtt atgaaactgg aggtggttagc tggaataaaa     1200
attgcacatc ctttcaccaa accctttgaa agtagttatt gttgtccaac cgaggatgac     1260
tatgaaatga ttcaagacaa atacggtagt cataaaactg agacagcact gaacgccctt     1320
aaactggtaa cagatgaaaa taaagaggaa gaaagtatta aagatgcacc aaaggcatat     1380
ttaagcacca tgtacatagg ccttgacttt aatattgaaa acaaaaagga aaaagttgac     1440
attcacattc cctgcactga atttgtgaat ttatgtcgaa gtttcaatga ggattatggt     1500
gaccacaaag tattcaatct agccctccgc ttcgtaaagg gttacgattt gccagatgaa     1560
gttttcgatg aaaatgaaaa gagaccatca aagaagagta aaaggaagaa ttagatgct      1620
agacatgaaa ccgtgaagag atctaaatca gatgctgctt caggtgacaa catcaatggc     1680
acaaccgcag ctgttgacgt aaactaa                                     1707

```

```

<210> 37
<211> 626
<212> PRT
<213> Saccharomyces cerevisiae

```

```

<400> 37

```

```

Met Asp His Asp Thr Glu Val Ile Val Lys Asp Phe Asn Ser Ile Leu
1              5              10              15

```

```

Glu Glu Leu Thr Phe Asn Ser Arg Pro Ile Ile Thr Thr Leu Thr Lys
                20              25              30

```

```

Leu Ala Glu Glu Asn Ile Ser Cys Ala Gln Tyr Phe Val Asp Ala Ile
        35              40              45

```

```

Glu Ser Arg Ile Glu Lys Cys Met Pro Lys Gln Lys Leu Tyr Ala Phe
50              55              60

```

```

Tyr Ala Leu Asp Ser Ile Cys Lys Asn Val Gly Ser Pro Tyr Thr Ile
65              70              75              80

```

```

Tyr Phe Ser Arg Asn Leu Phe Asn Leu Tyr Lys Arg Thr Tyr Leu Leu
            85              90              95

```

Val Asp Asn Thr Thr Arg Thr Lys Leu Ile Asn Met Phe Lys Leu Trp
 100 105 110

Leu Asn Pro Asn Asp Thr Gly Leu Pro Leu Phe Glu Gly Ser Ala Leu
 115 120 125

Glu Lys Ile Glu Gln Phe Leu Ile Lys Ala Ser Ala Leu His Gln Lys
 130 135 140

Asn Leu Gln Ala Met Leu Pro Thr Pro Thr Val Pro Leu Leu Leu Arg
 145 150 155 160

Asp Ile Asp Lys Leu Thr Cys Leu Thr Ser Glu Arg Leu Lys Asn Gln
 165 170 175

Pro Asn Asp Glu Lys Leu Lys Met Lys Leu Leu Val Leu Ser Gln Leu
 180 185 190

Lys Gln Glu Leu Lys Arg Glu Lys Leu Thr Leu Asn Ala Leu Lys Gln
 195 200 205

Val Gln Met Gln Leu Arg Gln Val Phe Ser Gln Asp Gln Gln Val Leu
 210 215 220

Gln Glu Arg Met Arg Tyr His Glu Leu Gln Gln Gln Gln Gln Gln
 225 230 235 240

Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Tyr His Glu
 245 250 255

Thr Lys Asp Met Val Gly Ser Tyr Thr Gln Asn Ser Asn Ser Ala Ile
 260 265 270

Pro Leu Phe Gly Asn Asn Ser Asp Thr Thr Asn Gln Gln Asn Ser Leu
 275 280 285

Ser Ser Ser Leu Phe Gly Asn Ile Ser Gly Val Glu Ser Phe Gln Glu
 290 295 300

Ile Glu Lys Lys Lys Ser Leu Asn Lys Ile Asn Asn Leu Tyr Ala Ser
 305 310 315 320

Leu Lys Ala Glu Gly Leu Ile Tyr Thr Pro Pro Lys Glu Ser Ile Val
 325 330 335

Thr Leu Tyr Lys Lys Leu Asn Gly His Ser Asn Tyr Ser Leu Asp Ser
 340 345 350

His Glu Lys Gln Leu Met Lys Asn Leu Pro Lys Ile Pro Leu Leu Asn
 355 360 365

Asp Ile Leu Ser Asp Cys Lys Ala Tyr Phe Ala Thr Val Asn Ile Asp
 370 375 380

Val Leu Asn Asn Pro Ser Leu Gln Leu Ser Glu Gln Thr Leu Leu Gln
 385 390 395 400

Glu Asn Pro Ile Val Gln Asn Asn Leu Ile His Leu Leu Tyr Arg Ser
 405 410 415

Lys Pro Asn Lys Cys Ser Val Cys Gly Lys Arg Phe Gly Asn Ser Glu
 420 425 430

Ser Glu Lys Leu Leu Gln Asn Glu His Leu Asp Trp His Phe Arg Ile
 435 440 445

Asn Thr Arg Ile Lys Gly Ser Gln Asn Thr Ala Asn Thr Gly Ile Ser
 450 455 460

Asn Ser Asn Leu Asn Thr Thr Thr Thr Arg Lys Asn Ile Gln Ser Arg
 465 470 475 480

Asn Trp Tyr Leu Ser Asp Ser Gln Trp Ala Ala Phe Lys Asp Asp Glu
 485 490 495

Ile Thr Ser Thr Lys His Lys Asn Asp Tyr Thr Asp Pro His Ala Asn
 500 505 510

Lys Asn Ile Asp Lys Ser Ala Leu Asn Ile His Ala Asp Glu Asn Asp
 515 520 525

Glu Gly Ser Val Asp Asn Thr Leu Gly Ser Asp Arg Ser Asn Glu Leu
 530 535 540

Glu Ile Arg Gly Lys Tyr Val Val Val Pro Glu Thr Ser Gln Asp Met
 545 550 555 560

Ala Phe Lys Cys Pro Ile Cys Lys Glu Thr Val Thr Gly Val Tyr Asp
 565 570 575

Glu Glu Ser Gly Glu Trp Val Trp Lys Asn Thr Ile Glu Val Asn Gly
 580 585 590

Lys Tyr Phe His Ser Thr Cys Tyr His Glu Thr Ser Gln Asn Ser Ser

595

600

605

Lys Ser Asn Ser Gly Lys Val Gly Leu Asp Asp Leu Lys Lys Leu Val
 610 615 620

Thr Lys
 625

<210> 38

<211> 1881

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 38

```

atggatcacg acacagaagt tatagtcaag gatttcaaca gtattcttga agaattaact      60
tttaactcga gaccattat taccactctc acgaagttag cagaagaaaa catttcgtgt      120
gctcaatatt ttgttgacgc cattgagagt agaattgaaa aatgtatgcc taagcagaag      180
ctatacgctt tttatgcctt ggactccatt tgtaagaatg ttggcagccc atacacaatc      240
tacttcagta gaaatttggt taacttgtac aaaagaacat acttattggt ggacaacacc      300
acaagaacaa aactcatcaa tatgtttaag ctttggttga atcctaata caccggcctg      360
cctctgtttg aaggctcagc actggagaaa attgaacagt ttttaataaa ggccagtgtc      420
ctccaccaga aaaacctgca agcaatgtta ccaactccga ctgttcccct cctattaaga      480
gacattgata aattaacttg tttaactagt gaaaggctaa agaatcaacc aaatgatgaa      540
aaattaaaaa tgaaactggt ggttttatct cagttgaaac aagaactgaa aagagaaaaa      600
ctaactctca atgctctgaa acaagttcaa atgcaactaa ggcaggtctt ctcacaggac      660
caacaagttt tgcaggaaag aatgaggtat cacgaacttc agcagcaaca gcaacaacag      720
cagcaacagc aacaacaaca acagcagcag cagcagcagt atcacgaaac taaggatatg      780
gtcggttctt acacacaaaa ctctaattca gcgattccat tgtttgaaa taattcagat      840
acaacgaacc aacagaattc tttatcttcc tctctttttg gtaatatctt tggcgttgaa      900
tcttttcaag aaattgagaa aaaaaagtcc ctaaacaaaa tcaataacct ctatgcatct      960
cttaaggctg aagggtcat atatactcca ccaaagaat ccatcgtaac gctttataaa     1020
aagctaaatg gacactcaa ctactctctt gattcacacg aaaaacaatt gatgaagaac     1080
ctgccaaaga tacccttatt aaatgatatt ctttctgact gcaaagctta ttttgccacg     1140
gttaacatcg acgttttgaa taacctatcc ctgcagctgt ccgaacaaac tcttcttcaa     1200
gaaaatccca tcgtccaaaa taatcttatt cacttattat accgatccaa gccaaataaa     1260
tgtagcgtgt gtggtaagag atttggcaac tcagaaagtg aaaaactact acaaaatgag     1320
catttggtgact ggcatttcag aataaataca agaattaagg gatcacaaaa cacagcaaata     1380

```

```

acaggcataa gcaactctaa cttaaataca actaccacaa ggaaaaatat tcaatcaaga 1440
aactgggtatt tgagtgactc tcaatgggca gccttcaaag atgatgaaat tacctccaca 1500
aagcataaaa atgattatac tgatccgcat gccataaaaa atatcgataa aagtgccta 1560
aatattcacg ctgatgaaaa tgacgaaggc tctgtagaca acactttagg aagtgataga 1620
agtaacgagt tggaataacg gggaaaaatat gttgtgttac cagaaacttc acaagatatg 1680
gccttcaaat gtccaatttg taaggaaacc gttactggcg tttatgatga agaactctgg 1740
gaatgggttt ggaaaaatac tattgaagtg aatggcaaat attccactc gacatgttat 1800
catgaaacgt cgcaaaattc aagcaaactc aatagtggca aggtcgggtt ggatgactta 1860
aagaaattgg tcacaaaata a 1881

```

```

<210> 39
<211> 563
<212> PRT
<213> Saccharomyces cerevisiae

```

```

<400> 39

```

```

Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Lys Gln
1           5           10          15

```

```

Val Asn Val Asn Thr Val Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
          20          25          30

```

```

Leu Leu Asp Lys Ile Tyr Glu Val Glu Gly Met Arg Trp Ala Gly Asn
          35          40          45

```

```

Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
          50          55          60

```

```

Lys Gly Met Ser Cys Ile Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
          65          70          75          80

```

```

Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
          85          90          95

```

```

His Val Val Gly Val Pro Ser Ile Ser Ala Gln Ala Lys Gln Leu Leu
          100          105          110

```

```

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
          115          120          125

```

```

Ser Ala Asn Ile Ser Glu Thr Thr Ala Met Ile Thr Asp Ile Ala Thr
          130          135          140

```

```

Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Val Thr Gln

```

145		150		155		160
Arg Pro Val Tyr	Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Asn Val					
	165		170		175	
Pro Ala Lys Leu Leu Gln Thr Pro Ile Asp Met Ser Leu Lys Pro Asn						
	180		185		190	
Asp Ala Glu Ser Glu Lys Glu Val Ile Asp Thr Ile Leu Ala Leu Val						
	195		200		205	
Lys Asp Ala Lys Asn Pro Val Ile Leu Ala Asp Ala Cys Cys Ser Arg						
	210		215		220	
His Asp Val Lys Ala Glu Thr Lys Lys Leu Ile Asp Leu Thr Gln Phe						
	225		230		235	240
Pro Ala Phe Val Thr Pro Met Gly Lys Gly Ser Ile Asp Glu Gln His						
	245		250			255
Pro Arg Tyr Gly Gly Val Tyr Val Gly Thr Leu Ser Lys Pro Glu Val						
	260		265		270	
Lys Glu Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Val Gly Ala Leu						
	275		280		285	
Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys						
	290		295		300	
Asn Ile Val Glu Phe His Ser Asp His Met Lys Ile Arg Asn Ala Thr						
	305		310		315	320
Phe Pro Gly Val Gln Met Lys Phe Val Leu Gln Lys Leu Leu Thr Thr						
	325		330			335
Ile Ala Asp Ala Ala Lys Gly Tyr Lys Pro Val Ala Val Pro Ala Arg						
	340		345		350	
Thr Pro Ala Asn Ala Ala Val Pro Ala Ser Thr Pro Leu Lys Gln Glu						
	355		360		365	
Trp Met Trp Asn Gln Leu Gly Asn Phe Leu Gln Glu Gly Asp Val Val						
	370		375		380	
Ile Ala Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Thr Phe						
	385		390		395	400

Pro Asn Asn Thr Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
 405 410 415

Phe Thr Thr Gly Ala Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
 420 425 430

Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
 435 440 445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
 450 455 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Lys Leu Ile
 465 470 475 480

His Gly Pro Lys Ala Gln Tyr Asn Glu Ile Gln Gly Trp Asp His Leu
 485 490 495

Ser Leu Leu Pro Thr Phe Gly Ala Lys Asp Tyr Glu Thr His Arg Val
 500 505 510

Ala Thr Thr Gly Glu Trp Asp Lys Leu Thr Gln Asp Lys Ser Phe Asn
 515 520 525

Asp Asn Ser Lys Ile Arg Met Ile Glu Ile Met Leu Pro Val Phe Asp
 530 535 540

Ala Pro Gln Asn Leu Val Glu Gln Ala Lys Leu Thr Ala Ala Thr Asn
 545 550 555 560

Ala Lys Gln

<210> 40

<211> 1692

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 40

atgtctgaaa ttactttggg taaatatttg ttcgaaagat taaagcaagt caacgttaac	60
accgttttgc gtttgccagg tgacttcaac ttgtccttgc tggacaagat ctacgaagtt	120
gaaggatatga gatgggctgg taacgccaac gaattgaacg ctgcttacgc cgctgatggg	180
tacgctcgta tcaaggggat gtcttgtatc atcaccacct tcggtgtcgg tgaattgtct	240
gctttgaacg gtattgccgg ttcttacgct gaacacgtcg gtgttttgca cgttgttggt	300
gtcccatcca tctctgctca agctaagcaa ttgtgtgtgc accacacctt gggtaacggg	360
gacttcactg ttttccacag aatgtctgcc aacatttctg aaaccactgc tatgatcact	420

```

gacattgcta cgcggccagc tgaaattgac agatgtatca gaaccactta cgtcacccaa   480
agaccagtct acttaggttt gccagctaac ttggctgact tgaacgtccc agctaagttg   540
ttgcaaaactc caattgacat gtctttgaag ccaaacgatg ctgaatccga aaaggaagtc   600
attgacacca tcttggcttt ggtcaaggat gctaagaacc cagttatctt ggctgatgct   660
tggtgttcca gacacgacgt caaggctgaa actaagaagt tgattgactt gactcaattc   720
ccagctttcg tcaccccaat gggtaagggt tccattgacg aacaacaccc aagatacggg   780
gggtgtttacg tcggtacctt gtccaagcca gaagttaagg aagccgttga atctgctgac   840
ttgattttgt ctgctgggtgc tttgttgtct gatttcaaca cgggttcttt ctcttactct   900
tacaagacca agaacattgt cgaattccac tccgaccaca tgaagatcag aaacgccact   960
ttcccagggtg tccaaatgaa attcgttttg caaaagttgt tgaccactat tgctgacgcc  1020
gctaagggtt acaagccagt tgctgtccca gctagaactc cagctaacgc tgctgtccca  1080
gcttctaccc cattgaagca agaattggatg tggaaccaat tgggtaactt cttgcaagaa  1140
gggtgatgttg tcattgctga aaccgggtacc tccgctttcg gtatcaacca aaccactttc  1200
ccaaacaaca cctacggtat ctctcaagtc ttatgggggt ccattgggtt caccactggg  1260
gctaccttgg gtgctgcttt cgctgctgaa gaaattgatc caaagaagag agttatctta  1320
ttcattgggtg acggttcttt gcaattgact gttcaagaaa tctccaccat gatcagatgg  1380
ggcttgaagc catacttggt cgtcttgaac aacgatgggt acaccattga aaagttgatt  1440
cacggtccaa aggtcaata caacgaaatt caaggttggg accacctatc cttgttgcca  1500
actttcggtg ctaaggacta tgaaaccac agagtcgcta ccaccgggtga atgggacaag  1560
ttgacccaag acaagtcttt caacgacaac tctaagatca gaatgattga aatcatgttg  1620
ccagtcttcg atgctccaca aaacttggtt gaacaagcta agttgactgc tgctaccaac  1680
gctaagcaat aa                                     1692

```

```

<210> 41
<211> 987
<212> PRT
<213> Saccharomyces cerevisiae

```

```

<400> 41

```

```

Met Gln Ser Gln Asp Ser Cys Tyr Gly Val Ala Phe Arg Ser Ile Ile
1           5           10           15

```

```

Thr Asn Asp Glu Ala Leu Phe Lys Lys Thr Ile His Phe Tyr His Thr
20           25           30

```

```

Leu Gly Phe Ala Thr Val Lys Asp Phe Asn Lys Phe Lys His Gly Glu
35           40           45

```

Asn Ser Leu Leu Ser Ser Gly Thr Ser Gln Asp Ser Leu Arg Glu Val
 50 55 60

Trp Leu Glu Ser Phe Lys Leu Ser Glu Val Asp Ala Ser Gly Phe Arg
 65 70 75 80

Ile Pro Gln Gln Glu Ala Thr Asn Lys Ala Gln Ser Gln Gly Ala Leu
 85 90 95

Leu Lys Ile Arg Leu Val Met Ser Ala Pro Ile Asp Glu Thr Phe Asp
 100 105 110

Thr Asn Glu Thr Ala Thr Ile Thr Tyr Phe Ser Thr Asp Leu Asn Lys
 115 120 125

Ile Val Glu Lys Phe Pro Lys Gln Ala Glu Lys Leu Ser Asp Thr Leu
 130 135 140

Val Phe Leu Lys Asp Pro Met Gly Asn Asn Ile Thr Phe Ser Gly Leu
 145 150 155 160

Ala Asn Ala Thr Asp Ser Ala Pro Thr Ser Lys Asp Ala Phe Leu Glu
 165 170 175

Ala Thr Ser Glu Asp Glu Ile Ile Ser Arg Ala Ser Ser Asp Ala Ser
 180 185 190

Asp Leu Leu Arg Gln Thr Leu Gly Ser Ser Gln Lys Lys Lys Lys Ile
 195 200 205

Ala Val Met Thr Ser Gly Gly Asp Ser Pro Gly Met Asn Ala Ala Val
 210 215 220

Arg Ala Val Val Arg Thr Gly Ile His Phe Gly Cys Asp Val Phe Ala
 225 230 235 240

Val Tyr Glu Gly Tyr Glu Gly Leu Leu Arg Gly Gly Lys Tyr Leu Lys
 245 250 255

Lys Met Ala Trp Glu Asp Val Arg Gly Trp Leu Ser Glu Gly Gly Thr
 260 265 270

Leu Ile Gly Thr Ala Arg Ser Met Glu Phe Arg Lys Arg Glu Gly Arg
 275 280 285

Arg Gln Ala Ala Gly Asn Leu Ile Ser Gln Gly Ile Asp Ala Leu Val

290		295		300
Val Cys Gly Gly Asp Gly Ser Leu Thr Gly Ala Asp Leu Phe Arg His				
305		310	315	320
Glu Trp Pro Ser Leu Val Asp Glu Leu Val Ala Glu Gly Arg Phe Thr				
	325		330	335
Lys Glu Glu Val Ala Pro Tyr Lys Asn Leu Ser Ile Val Gly Leu Val				
	340		345	350
Gly Ser Ile Asp Asn Asp Met Ser Gly Thr Asp Ser Thr Ile Gly Ala				
	355		360	365
Tyr Ser Ala Leu Glu Arg Ile Cys Glu Met Val Asp Tyr Ile Asp Ala				
	370		375	380
Thr Ala Lys Ser His Ser Arg Ala Phe Val Val Glu Val Met Gly Arg				
	385		390	395
His Cys Gly Trp Leu Ala Leu Met Ala Gly Ile Ala Thr Gly Ala Asp				
	405		410	415
Tyr Ile Phe Ile Pro Glu Arg Ala Val Pro His Gly Lys Trp Gln Asp				
	420		425	430
Glu Leu Lys Glu Val Cys Gln Arg His Arg Ser Lys Gly Arg Arg Asn				
	435		440	445
Asn Thr Ile Ile Val Ala Glu Gly Ala Leu Asp Asp Gln Leu Asn Pro				
	450		455	460
Val Thr Ala Asn Asp Val Lys Asp Ala Leu Ile Glu Leu Gly Leu Asp				
	465		470	475
Thr Lys Val Thr Ile Leu Gly His Val Gln Arg Gly Gly Thr Ala Val				
	485		490	495
Ala His Asp Arg Trp Leu Ala Thr Leu Gln Gly Val Asp Ala Val Lys				
	500		505	510
Ala Val Leu Glu Phe Thr Pro Glu Thr Pro Ser Pro Leu Ile Gly Ile				
	515		520	525
Leu Glu Asn Lys Ile Ile Arg Met Pro Leu Val Glu Ser Val Lys Leu				
	530		535	540

Thr Lys Ser Val Ala Thr Ala Ile Glu Asn Lys Asp Phe Asp Lys Ala
545 550 555 560

Ile Ser Leu Arg Asp Thr Glu Phe Ile Glu Leu Tyr Glu Asn Phe Leu
565 570 575

Ser Thr Thr Val Lys Asp Asp Gly Ser Glu Leu Leu Pro Val Ser Asp
580 585 590

Arg Leu Asn Ile Gly Ile Val His Val Gly Ala Pro Ser Ala Ala Leu
595 600 605

Asn Ala Ala Thr Arg Ala Ala Thr Leu Tyr Cys Leu Ser His Gly His
610 615 620

Lys Pro Tyr Ala Ile Met Asn Gly Phe Ser Gly Leu Ile Gln Thr Gly
625 630 635 640

Glu Val Lys Glu Leu Ser Trp Ile Asp Val Glu Asn Trp His Asn Leu
645 650 655

Gly Gly Ser Glu Ile Gly Thr Asn Arg Ser Val Ala Ser Glu Asp Leu
660 665 670

Gly Thr Ile Ala Tyr Tyr Phe Gln Lys Asn Lys Leu Asp Gly Leu Ile
675 680 685

Ile Leu Gly Gly Phe Glu Gly Phe Arg Ser Leu Lys Gln Leu Arg Asp
690 695 700

Gly Arg Thr Gln His Pro Ile Phe Asn Ile Pro Met Cys Leu Ile Pro
705 710 715 720

Ala Thr Val Ser Asn Asn Val Pro Gly Thr Glu Tyr Ser Leu Gly Val
725 730 735

Asp Thr Cys Leu Asn Ala Leu Val Asn Tyr Thr Asp Asp Ile Lys Gln
740 745 750

Ser Ala Ser Ala Thr Arg Arg Arg Val Phe Val Cys Glu Val Gln Gly
755 760 765

Gly His Ser Gly Tyr Ile Ala Ser Phe Thr Gly Leu Ile Thr Gly Ala
770 775 780

Val Ser Val Tyr Thr Pro Glu Lys Lys Ile Asp Leu Ala Ser Ile Arg
785 790 795 800

Glu Asp Ile Thr Leu Leu Lys Glu Asn Phe Arg His Asp Lys Gly Glu
805 810 815

Asn Arg Asn Gly Lys Leu Leu Val Arg Asn Glu Gln Ala Ser Ser Val
820 825 830

Tyr Ser Thr Gln Leu Leu Ala Asp Ile Ile Ser Glu Ala Ser Lys Gly
835 840 845

Lys Phe Gly Val Arg Thr Ala Ile Pro Gly His Val Gln Gln Gly Gly
850 855 860

Val Pro Ser Ser Lys Asp Arg Val Thr Ala Ser Arg Phe Ala Val Lys
865 870 875 880

Cys Ile Lys Phe Ile Glu Gln Trp Asn Lys Lys Asn Glu Ala Ser Pro
885 890 895

Asn Thr Asp Ala Lys Val Leu Arg Phe Lys Phe Asp Thr His Gly Glu
900 905 910

Lys Val Pro Thr Val Glu His Glu Asp Asp Ser Ala Ala Val Ile Cys
915 920 925

Val Asn Gly Ser His Val Ser Phe Lys Pro Ile Ala Asn Leu Trp Glu
930 935 940

Asn Glu Thr Asn Val Glu Leu Arg Lys Gly Phe Glu Val His Trp Ala
945 950 955 960

Glu Tyr Asn Lys Ile Gly Asp Ile Leu Ser Gly Arg Leu Lys Leu Arg
965 970 975

Ala Glu Val Ala Ala Leu Ala Ala Glu Asn Lys
980 985

<210> 42

<211> 2000

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 42

gatcacagtt ttctgaggat agcactgatc tatacaataa aaggagctat ttctatcata 60

atgaaagaga tgtatttgta ttcatTTTCg cacttgccac cattttacta attgtttgct 120

catgttatgt tacgcctcta tatcgtatgc atcacaagat gggaaacttg aattggtcct 180

cggcgtaga ttgcgccttc attggtgcct tctcaattga atttatcgtg aaaacagtag 240

```

ctgacggatt tatatattct ccaaagctt acctgaggaa tccatggaac tttattgatt 300
tttgtgtcct aatctcaatg tggattaatt taattgcata cctaaaaaac aatggaaatt 360
tgtctaggat tttcaaggga ttgacagccc tgagggccct cagatgcctc acgatcagta 420
acacagctcg tcaaacattt aacctagtta tgtttgatgg tttaaataaa atttttgaag 480
ctgggttgat ttcactcagt ttgctatttc catttacagt ttggggctta agcattttta 540
aaggccgttt aggtacttgc aatgacggaa gtttgggccc tgcagattgt tacaatgaat 600
attcaaatc cgtttttcaa tgggatatca tgtctccaag ggtttaccag caaccatata 660
ttcatttgga ctctttcgca agcgctttta gttcattata ccaaatcatt tctttggaag 720
gatgggttga tttgttgga aatatgatga atagtccagg aataggatca cccgctacgg 780
taatgggttc agcaggggaat gctttattcc tcgttctgtt taatttttta agtatgggtt 840
tcacctgaa cttgtttgtt tcattcattg ttaacaacca agcaaggaca acaggaagcg 900
cttactttac cattgaggaa aaggcgtggc tggaaatcca gaaactttta tctcaggcca 960
agccaaaagc tatcccaaat ttaattgagt tatcaagagt taggcaattt ttctatcaac 1020
ttgcagtgga gaaaaaaaaat ttctactacg catcgtttct tcaggtagta ctttatttgc 1080
acataatcat gctcctgagt cgaagctaca atccaggaaa cttgataggt tatcaagggtg 1140
tttattttat gttttccact agtgtttttt taattcaaga ggcacttcac atgtgcgggtg 1200
aaggaccaag attatatttt aggcaaaaat ggaacagcat acgactcagt atcataatta 1260
tagcctttat tatgaacgct gtagcattcc acgttccagc ctctcactat tggttccaca 1320
atataaaggg gtttttcctg ttagtgatat ttttgtttat tattcctcaa aatgacacac 1380
taactgaact attagaaacc gcaatggcaa gcttaccgcc tattctatca ttgacctaca 1440
cttggggggt tttattttta gtatatgcta ttgctttgaa tcaaactctt gcctaacaa 1500
ggttaggagg taatacgacc gataacataa attttagaac tgtaatcaaa tccatgattg 1560
ttctgtttag atgtagtttt ggtgagggtt ggaattatat catggccgac ctaactgtgt 1620
cagaacctta ttgctcctct gatgataatt caacctatac ggactgtgga tcagagacat 1680
atgcctattt gttattaatg tcgtggaata ttatttccat gtatattttt gtgaatatgt 1740
ttgtttcgtt gattatttgt aatttcagtt atgtttaccg tagcggtgga tctcgctctg 1800
gcatcaacag atcggagata aaaaaataca ttgaagcttg gtccaaattt gatactgatg 1860
gaactggcga gcttgagctg tcctacctcc caagaataat gcattcattt gacggctctc 1920
tttcatttaa aatttgggaa ggtagattga caataaaaag tctagtcgag aactacatgg 1980
aggttaaccc agatgatcca 2000

```

<210> 43

<211> 465

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 43

Met Asp Gly His Asn Gln Asn Gln Tyr Gln Asn Gln Asn Gln Ile Gln
 1 5 10 15

Gln Ser Gln Gln Pro Pro Leu Lys Lys Tyr Val Thr Gln Arg Arg Ser
 20 25 30

Val Asp Val Ser Ser Pro Tyr Ile Asn Leu Tyr Tyr Asn Arg Arg His
 35 40 45

Gly Leu Pro Asn Leu Val Val Glu Pro Glu Thr Ser Tyr Thr Ile Asp
 50 55 60

Ile Met Pro Pro Asn Ala Tyr Arg Gly Arg Asp Arg Val Ile Asn Leu
 65 70 75 80

Pro Ser Lys Phe Thr His Leu Ser Ser Asn Lys Val Lys His Val Ile
 85 90 95

Pro Ala Ile Gln Trp Thr Pro Glu Gly Arg Arg Leu Val Val Ala Thr
 100 105 110

Tyr Ser Gly Glu Phe Ser Leu Trp Asn Ala Ser Ser Phe Thr Phe Glu
 115 120 125

Thr Leu Met Gln Ala His Asp Ser Ala Val Thr Thr Met Lys Tyr Ser
 130 135 140

His Asp Ser Asp Trp Met Ile Ser Gly Asp Ala Asp Gly Met Ile Lys
 145 150 155 160

Ile Trp Gln Pro Asn Phe Ser Met Val Lys Glu Ile Asp Ala Ala His
 165 170 175

Thr Glu Ser Ile Arg Asp Met Ala Phe Ser Ser Asn Asp Ser Lys Phe
 180 185 190

Val Thr Cys Ser Asp Asp Asn Ile Leu Lys Ile Trp Asn Phe Ser Asn
 195 200 205

Gly Lys Gln Glu Arg Val Leu Ser Gly His His Trp Asp Val Lys Ser
 210 215 220

Cys Asp Trp His Pro Glu Met Gly Leu Ile Ala Ser Ala Ser Lys Asp
 225 230 235 240

Asn Leu Val Lys Leu Trp Asp Pro Arg Ser Gly Asn Cys Ile Ser Ser
 245 250 255

Ile Leu Lys Phe Lys His Thr Val Leu Lys Thr Arg Phe Gln Pro Thr
 260 265 270

Lys Gly Asn Leu Leu Met Ala Ile Ser Lys Asp Lys Ser Cys Arg Val
 275 280 285

Phe Asp Ile Arg Tyr Ser Met Lys Glu Leu Met Cys Val Arg Asp Glu
 290 295 300

Thr Asp Tyr Met Thr Leu Glu Trp His Pro Ile Asn Glu Ser Met Phe
 305 310 315 320

Thr Leu Ala Cys Tyr Asp Gly Ser Leu Lys His Phe Asp Leu Leu Gln
 325 330 335

Asn Leu Asn Glu Pro Ile Leu Thr Ile Pro Tyr Ala His Asp Lys Cys
 340 345 350

Ile Thr Ser Leu Ser Tyr Asn Pro Val Gly His Ile Phe Ala Thr Ala
 355 360 365

Ala Lys Asp Arg Thr Ile Arg Phe Trp Thr Arg Ala Arg Pro Ile Asp
 370 375 380

Pro Asn Ala Tyr Asp Asp Pro Thr Tyr Asn Asn Lys Lys Ile Asn Gly
 385 390 395 400

Trp Phe Phe Gly Ile Asn Asn Asp Ile Asn Ala Val Arg Glu Lys Ser
 405 410 415

Glu Phe Gly Ala Ala Pro Pro Pro Pro Ala Thr Leu Glu Pro His Ala
 420 425 430

Leu Pro Asn Met Asn Gly Phe Ile Asn Lys Lys Pro Arg Gln Glu Ile
 435 440 445

Pro Gly Ile Asp Ser Asn Ile Lys Ser Ser Thr Leu Pro Gly Leu Ser
 450 455 460

Ile
 465

<211> 1398

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 44

```

atggacgggc ataatcagaa ccagtaccag aatcagaatc aaattcagca gtcacaacaa      60
ccacctctga agaaatatgt gacacagaga cgatcggtag acgtgagttc tccgtatatt      120
aacctatatt acaacagacg acatggatta ccgaacctag tggtagagcc agaaacttca      180
tacacaatag atataatgcc gcctaatagcc tacagaggtc gagatcgagt cataaatttg      240
cccagcaaat ttacgcattt aagctcgaat aaagtgaaac atgtgatacc cgccatccaa      300
tggaccccag aaggcagaag acttgtagta gcaacttata gcggagaatt ttctctatgg      360
aacgcatctt cgtttacgtt tgaaaccctc atgcaagcac atgattcggc tgtgactaca      420
atgaaatatt ctacagacag tgattggatg attagtgggg acgcagatgg aatgattaag      480
atatggcaac caaactttag tatggttaaa gaaattgatg ctgcacacac ggaaagtatt      540
agggatatgg cattcagtag taacgactcg aaatttgtaa cctgttcaga tgataatatc      600
ttgaagattt ggaacttcag caatggcaaa caagaaagag tattatcggg gcaccactgg      660
gacgtgaaaa gttgtgattg gcatcccgag atgggggttaa ttgcctccgc atctaaggac      720
aatttagtaa aactgtggga tccgcgttcc ggaaactgca tatcatcaat tctgaagttc      780
aaacatactg tcttgaagac gcgatttcaa cccaccaagg gaaatctatt aatggcgatt      840
tcaaaagata aatcgtgtcg tgtgttcgac attagataca gtatgaagga actaatgtgt      900
gtaagagatg aaacggatta tatgacacta gaatggcacc caataaatga atctatgttc      960
actctggcgt gctatgacgg ctgcgctaaag cattttgatc ttttacaaaa tctaaacgaa    1020
cctattttga ctattccata cgcccatgat aagtgtatca cctcgttatc atataatcca    1080
gtaggacata tatttgccac agcagctaag gatagaacca taagattttg gactagagcg    1140
aggccaattg atcctaattg atatgatgat ccaacttata acaataagaa gataaatgga    1200
tggttttttg gcatcaataa cgacatcaat gctgtgagag agaagagtga attcgggtgcg    1260
gccccgccac caccagctac acttgaacca catgcgttac caaacatgaa tggattcatc    1320
aacaagaaac cgcgtcaaga aatcccaggc atcgattcaa acatcaagag ttccacattg    1380
ccaggtttaa gcatataa                                     1398

```

<210> 45

<211> 785

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 45

```

Met Ser Ser Ala Glu Met Glu Gln Leu Leu Gln Ala Lys Thr Leu Ala
1           5           10           15

```

Met His Asn Asn Pro Thr Glu Met Leu Pro Lys Val Leu Glu Thr Thr
 20 25 30

Ala Ser Met Tyr His Asn Gly Asn Leu Ser Lys Leu Lys Leu Pro Leu
 35 40 45

Ala Lys Phe Phe Thr Gln Leu Val Leu Asp Val Val Ser Met Asp Ser
 50 55 60

Pro Ile Ala Asn Thr Glu Arg Pro Phe Ile Ala Ala Gln Tyr Leu Pro
 65 70 75 80

Leu Leu Leu Ala Met Ala Gln Ser Thr Ala Asp Val Leu Val Tyr Lys
 85 90 95

Asn Ile Val Leu Ile Met Cys Ala Ser Tyr Pro Leu Val Leu Asp Leu
 100 105 110

Val Ala Lys Thr Ser Asn Gln Glu Met Phe Asp Gln Leu Cys Met Leu
 115 120 125

Lys Lys Phe Val Leu Ser His Trp Arg Thr Ala Tyr Pro Leu Arg Ala
 130 135 140

Thr Val Asp Asp Glu Thr Asp Val Glu Gln Trp Leu Ala Gln Ile Asp
 145 150 155 160

Gln Asn Ile Gly Val Lys Leu Ala Thr Ile Lys Phe Ile Ser Glu Val
 165 170 175

Val Leu Ser Gln Thr Lys Ser Pro Ser Gly Asn Glu Ile Asn Ser Ser
 180 185 190

Thr Ile Pro Asp Asn His Pro Val Leu Asn Lys Pro Ala Leu Glu Ser
 195 200 205

Glu Ala Lys Arg Leu Leu Asp Met Leu Leu Asn Tyr Leu Ile Glu Glu
 210 215 220

Gln Tyr Met Val Ser Ser Val Phe Ile Gly Ile Ile Asn Ser Leu Ser
 225 230 235 240

Phe Val Ile Lys Arg Arg Pro Gln Thr Thr Ile Arg Ile Leu Ser Gly
 245 250 255

Leu Leu Arg Phe Asn Val Asp Ala Lys Phe Pro Leu Glu Gly Lys Ser

260	265	270
Asp Leu Asn Tyr Lys Leu Ser Lys Arg Phe Val Glu Arg Ala Tyr Lys		
275	280	285
Asn Phe Val Gln Phe Gly Leu Lys Asn Gln Ile Ile Thr Lys Ser Leu		
290	295	300
Ser Ser Gly Ser Gly Ser Ser Ile Tyr Ser Lys Leu Thr Lys Ile Ser		
305	310	315
Gln Thr Leu His Val Ile Gly Glu Glu Thr Lys Ser Lys Gly Ile Leu		
325	330	335
Asn Phe Asp Pro Ser Lys Gly Asn Ser Lys Lys Thr Leu Ser Arg Gln		
340	345	350
Asp Lys Leu Lys Tyr Ile Ser Leu Trp Lys Arg Gln Leu Ser Ala Leu		
355	360	365
Leu Ser Thr Leu Gly Val Ser Thr Lys Thr Pro Thr Pro Val Ser Ala		
370	375	380
Pro Ala Thr Gly Ser Ser Thr Glu Asn Met Leu Asp Gln Leu Lys Ile		
385	390	395
Leu Gln Lys Tyr Thr Leu Asn Lys Ala Ser His Gln Gly Asn Thr Phe		
405	410	415
Phe Asn Asn Ser Pro Lys Pro Ile Ser Asn Thr Tyr Ser Ser Val Tyr		
420	425	430
Ser Leu Met Asn Ser Ser Asn Ser Asn Gln Asp Val Thr Gln Leu Pro		
435	440	445
Asn Asp Ile Leu Ile Lys Leu Ser Thr Glu Ala Ile Leu Gln Met Asp		
450	455	460
Ser Thr Lys Leu Ile Thr Gly Leu Ser Ile Val Ala Ser Arg Tyr Thr		
465	470	475
Asp Leu Met Asn Thr Tyr Ile Asn Ser Val Pro Ser Ser Ser Ser Ser		
485	490	495
Lys Arg Lys Ser Asp Asp Asp Asp Asp Gly Asn Asp Asn Glu Glu Val		
500	505	510

Gly Asn Asp Gly Pro Thr Ala Asn Ser Lys Lys Ile Lys Met Glu Thr
 515 520 525

Glu Pro Leu Ala Glu Glu Pro Glu Glu Pro Glu Asp Asp Arg Met
 530 535 540

Gln Lys Met Leu Gln Glu Glu Glu Ser Ala Gln Glu Ile Ser Gly Asp
 545 550 555 560

Ala Asn Lys Ser Thr Ser Ala Ile Lys Glu Ile Ala Pro Pro Phe Glu
 565 570 575

Pro Asp Ser Leu Thr Gln Asp Glu Lys Leu Lys Tyr Leu Ser Lys Leu
 580 585 590

Thr Lys Lys Leu Phe Glu Leu Ser Gly Arg Gln Asp Thr Thr Arg Ala
 595 600 605

Lys Ser Ser Ser Ser Ser Ser Ile Leu Leu Asp Asp Asp Asp Ser Ser
 610 615 620

Ser Trp Leu His Val Leu Ile Arg Leu Val Thr Arg Gly Ile Glu Ala
 625 630 635 640

Gln Glu Ala Ser Asp Leu Ile Arg Glu Glu Leu Leu Gly Phe Phe Ile
 645 650 655

Gln Asp Phe Glu Gln Arg Val Ser Leu Ile Ile Glu Trp Leu Asn Glu
 660 665 670

Glu Trp Phe Phe Gln Thr Ser Leu His Gln Asp Pro Ser Asn Tyr Lys
 675 680 685

Lys Trp Ser Leu Arg Val Leu Glu Ser Leu Gly Pro Phe Leu Glu Asn
 690 695 700

Lys His Arg Arg Phe Phe Ile Arg Leu Met Ser Glu Leu Pro Ser Leu
 705 710 715 720

Gln Ser Asp His Leu Glu Ala Leu Lys Pro Ile Cys Leu Asp Pro Ala
 725 730 735

Arg Ser Ser Leu Gly Phe Gln Thr Leu Lys Phe Leu Ile Met Phe Arg
 740 745 750

Pro Pro Val Gln Asp Thr Val Arg Asp Leu Leu His Gln Leu Lys Gln
 755 760 765

Glu Asp Glu Gly Leu His Lys Gln Cys Asp Ser Leu Leu Asp Arg Leu
 770 775 780

Lys
 785

<210> 46
 <211> 2000
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 46
 atgtcatctg cagagatgga acaattgtta caggccaaga cactggccat gcacaacaat 60
 ccaacggaga tgctgccaa ggtgctcgaa actacggcat ccatgtacca caacggtaat 120
 ctcagcaagc tgaagttgcc tttggccaag ttttttacac agttagttct agacgtggtg 180
 tcgatggact ctccaattgc gaatactgag agaccgttta ttgctgctca atatctgcc 240
 ctacttcttg ctatggcgca atccaccgcg gacgtactag tgtacaagaa tatcgtgctt 300
 attatgtgcg cttcataccc gctgggtgtg gatctgggtg ctaagacatc aaaccaggaa 360
 atgtttgatc agttgtgtat gctgaagaag ttcgtgctct cgcactggag aactgcatat 420
 cctttgcgtg ccaccgttga cgatgaaacg gatgtcgaa aatggctggc gcagattgac 480
 caaaatatcg gcgtgaaatt agcgaccatc aagttcatat ctgaggtcgt gctgtcgcaa 540
 actaaatcac ccagcggcaa cgagattaat tcatctacca tcccggataa ccaccctgtg 600
 ttgaacaaac cggctttgga gagcgaggct aagaggcttc ttgatatgtt gctaaactac 660
 ctaattgagg aacagtacat ggtctcgccc gttttcattg gtatcatcaa ttctttatcc 720
 ttcgtcatca aaagaaggcc gcagacaaca ataagaattc tttccgggct gttgcgtttc 780
 aacgtcgacg ccaagtttcc cctagagggc aagtctgact tgaactacaa actatccaag 840
 agatttggtg aaagggcgta caagaacttt gtgcaatttg ggctaaaaaa tcaaatcatt 900
 acaaaatccc tctcatccgg atcagggtea tcgatctact ccaagctgac caagatttct 960
 caaactttac acgttattgg cgaagagacc aagagcaagg gaattttgaa cttcgaccct 1020
 tccaagggca atagcaagaa aacgttgtcc aggcaggaca aactaaaata catctcacta 1080
 tggaaaaggc aattatccgc gttattgtct actctagggg tgtccacaaa gacccccacg 1140
 cctgtgtccg cacctgcaac gggctcttca accgaaaaca tgcttgatca actgaagata 1200
 ttgcaaaaat acaccctcaa caaggcttca caccagggca atactttttt caacaactca 1260
 cccaaaccaa tcagcaacac ctactcatct gtgtactcat tgatgaacag ttcgaactcc 1320
 aaccaggatg tgaccagct acccaatgac atacttatca agctgtccac agaggccatc 1380
 ttgcaaatgg acagcacgaa actgatcacc ggattgtcta tcgttgcttc gaggtacacg 1440

gatttaatga atacgtacat caattctgta ccgtcctcgt catcatcaaa gaggaaatcc 1500
 gacgatgatg acgacggcaa cgacaatgaa gaagttggaa acgatggccc aacggctaata 1560
 agcaagaaaa tcaaaatgga aacagaacca ctagcggagg aaccagagga gcccgaagac 1620
 gatgaccgaa tgcagaagat gcttcaagaa gaggaaagcg cccaagaaat ctcaggagat 1680
 gccacaacaa caacttctgc cattaaggag atcgcacccc cctttgaacc tgactcattg 1740
 acgcaggatg aaaaactaaa gtacctctca aagctgacca agaaactgtt tgaattatcc 1800
 ggtcgccagg atactaccgg ggccaaatct tcgtcttcct cctccatatt actggacgat 1860
 gacgactcct cgtcatgggt acacgtctta atcagattgg ttacgagagg aatcgaagca 1920
 caagaggcca gtgacctgat tcgtgaagaa ctgcttggtt tcttcatcca ggatttcgag 1980
 caacgtgtca gtctgatcat 2000

<210> 47
 <211> 533
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 47

Met Ser Ala Pro Val Pro Gln Leu Val Asn Ile Ser His Ala Leu Gln
 1 5 10 15

Ala Ser Thr Ile Gln Gln Ile Arg Leu Asp Met Val Asp Phe Asn Lys
 20 25 30

Asp Cys Lys Leu Ser Ser Ile Gln Leu Ala Arg Ile Asp Lys Tyr Ile
 35 40 45

Asp Ser Leu Gln Ala Ala Leu Asn Gln Phe Thr Lys Asp Asn Leu His
 50 55 60

Ile Glu Arg Lys Glu Lys Asn Val Thr Glu Ala Asp Ile Gln Leu Tyr
 65 70 75 80

Ser Gly Leu Lys Ser Met Tyr Leu Asp Tyr Leu Asn Gln Leu Ile Lys
 85 90 95

Leu Lys His Glu Lys Gln His His Ser Thr Pro Pro Ile Ala Asn Asp
 100 105 110

Val Ser Leu Asp Phe Phe Val Asn Gln Leu Pro Lys Phe Ser Pro Glu
 115 120 125

Glu Arg Lys Asn Tyr Ile Asp Asn Leu Ile Leu Asn Lys Asn Ser His
 130 135 140

Asn Arg Leu Ser Lys Met Asp Gly Leu Val Asp Ala Val Ile Asn Leu
 145 150 155 160

Cys Val Leu Asp Thr Ser Val Ala Glu Asn Val Arg Ser Tyr Met Lys
 165 170 175

Leu Leu Asp Thr Leu Gly Phe Gln Lys Gly Ser Asn Ser Thr Gly Thr
 180 185 190

Lys Ala Asn Leu Lys Lys Lys Leu Ala Ser Ser Lys Ala Lys Ile Lys
 195 200 205

Asp Ser Glu Lys Glu Lys Glu Lys Glu Lys Asp Lys Ser Lys Val Lys
 210 215 220

Met Lys Thr Lys Leu Lys Pro Ser Pro Leu Leu Asn Asn Asp Asp Lys
 225 230 235 240

Asn Ser Ser Pro Ser Pro Thr Ala Ser Thr Ser Ser Met Lys Lys Leu
 245 250 255

Lys Ser Gly Leu Phe Asn Lys Asn Glu Ala Lys Ser Thr Glu Ser Leu
 260 265 270

Pro Thr Ser Ser Lys Lys Lys Leu Ser Phe Ser Lys Tyr Leu Asn Lys
 275 280 285

Asp Asp Ala Asp Met Thr Lys Leu Gly Thr Lys Arg Ser Ile Asp Val
 290 295 300

Asp Phe Lys Val Asn Pro Glu Ala Ser Thr Val Ala Ser Asn Ile Ile
 305 310 315 320

Ser Ser Ser Thr Ser Gly Ser Ser Thr Thr Thr Val Ala Thr Pro Ala
 325 330 335

Ser Ser Glu Glu Pro Leu Lys Lys Lys Thr Lys Ile Ser Val Gln Asp
 340 345 350

Ser Asn Val Gln Ser Ile Leu Arg Asn Gly Lys Pro Lys Lys Ala Arg
 355 360 365

Ile Ser Ser Ile Lys Phe Leu Asp Asp Ser Gln Leu Ile Lys Val Tyr
 370 375 380

Gly Asp Asp Leu Pro Asn Gln Gly Leu Gln Val Ser Pro Thr Gln Leu
 385 390 395 400

Lys Lys Ile Leu Lys Pro Phe Lys Glu Gly Glu Pro Lys Glu Ile Ile
 405 410 415

Leu Phe Glu Asp Met Ser Ile Lys Leu Lys Pro Leu Asp Leu Met Phe
 420 425 430

Leu Lys Asn Thr Asn Ser Asp Asp Tyr Met Asp Ile Ser Glu Thr Lys
 435 440 445

Gly Gly Pro Ile His Cys Glu Thr Arg Thr Pro Leu Ile Tyr Arg Lys
 450 455 460

Asn Phe Asn His Phe Asn Pro Asp Leu Asn Lys Arg Pro Pro Arg Glu
 465 470 475 480

Pro Ile Glu Phe Asp Leu Asn Gly Asn Thr Asn Ser Thr Pro Thr Ile
 485 490 495

Ala Lys Ala Phe Gly Lys Asn Ser Leu Leu Leu Arg Lys Asp Arg Gly
 500 505 510

Gly Leu Pro Tyr Lys His Val Pro Ile Val Lys Arg Asn Lys Tyr Pro
 515 520 525

Pro Arg Pro Val His
 530

<210> 48
 <211> 1602
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 48
 atgtcagcac ctgttcctca attagtaaatt atttctcatg ccttgcaagc ttccactatc 60
 cagcaaattc gtttggatat ggtagatttt aataaagatt gcaaattatc ttccattcaa 120
 ctggcaagaa ttgacaagta tatcgattct ttgcaagcgg ctttgaatca gtttaccaaa 180
 gataacttgc acatagaacg gaaggaaaag aacgtgactg aagcagatat acagctgtat 240
 tctggcttaa aatcaatgta tctagattat ttgaaccagc taattaaact aaagcatgag 300
 aaacaacacc actctacacc gcccatcgct aatgatgtct ccctggactt cttcggtaac 360
 caattgccta agttttctcc cgaggaaagg aaaaattaca ttgacaattt aatcctgaat 420
 aaaaatagcc ataatcgatt atctaaaatg gatgggtctgg tagatgcagt catcaattta 480
 tgcgtcttgg atacttccgt tgctgaaaat gtacgatcct atatgaaatt attagatata 540
 ttgggtttcc aaaagggttc gaatagtact ggtacaaagg ccaacctcaa aaagaaattg 600

gctagttcga aagcaaagat aaaggattca gaaaaagaaa aggagaagga gaaggataaa 660
 tcaaaagtca agatgaaaac taaattaaaa ctttctcctt tgctcaataa cgatgacaaa 720
 aattcttctc catcgctac tgcattctacc tcttcaatga agaaattgaa gtcgggttta 780
 ttcaataaaa atgaagctaa gtctacagaa tctctaccta cttcttccaa gaaaaaacta 840
 tcattttcta aatatctgaa caaggatgac gcagatatga ccaagcttgg gactaaacgg 900
 tcaatagatg tggatttcaa agtcaacccc gaagcatcca cgggtggcttc taatatcata 960
 tcttcgtcaa cgtcaggatc gtcaaccaca acggtagcga ctctgcttc ttcagaagag 1020
 cccttaaaaa aaaaaaccaa aatatccgtg caagactcta atgtacaatc gattttgaga 1080
 aatggtaaac cgaaaaaagc acgcataagt agcatcaaat ttttggatga ttcccaacta 1140
 ataaaagttt acggtgacga tctaccgaac caagggtac aagtttctcc tactcaattg 1200
 aaaaaattc tgaaaccatt caaggagggg gaaccgaagg aaattatatt gttcgaggat 1260
 atgtcaatca aattaaaacc tcttgatttg atgtttctga agaacacaaa cagtgatgac 1320
 tatatggata tatccgagac taaaggtggc ccaatacatt gtgaaacaag gaccccgttg 1380
 atctatagaa aaaatttcaa tcatttcaac cgggacttga ataaaaggcc gccaaagagaa 1440
 cccatagaat tcgacttaaa tggaaatacg aactcaaccc cgactatagc aaaggctttc 1500
 ggtaaaaata gtttattact aaggaaggac agaggtgggt tgccatacaa gcatgtcccc 1560
 atagtaaaaa gaaataaata tcctccaaga ccagtacact aa 1602

<210> 49

<211> 677

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 49

Met Ser Ser Ser Thr Thr Pro Asp Leu Leu Tyr Pro Ser Ala Asp Lys
 1 5 10 15

Val Ala Glu Pro Ser Asp Asn Ile His Gly Asp Glu Leu Arg Leu Arg
 20 25 30

Glu Arg Ile Lys Asp Asn Pro Thr Asn Ile Leu Ser Tyr Phe Gln Leu
 35 40 45

Ile Gln Tyr Leu Glu Thr Gln Glu Ser Tyr Ala Lys Val Arg Glu Val
 50 55 60

Tyr Glu Gln Phe His Asn Thr Phe Pro Phe Tyr Ser Pro Ala Trp Thr
 65 70 75 80

Leu Gln Leu Lys Gly Glu Leu Ala Arg Asp Glu Phe Glu Thr Val Glu
 85 90 95

Lys Ile Leu Ala Gln Cys Leu Ser Gly Lys Leu Glu Asn Asn Asp Leu
 100 105 110

Ser Leu Trp Ser Thr Tyr Leu Asp Tyr Ile Arg Arg Lys Asn Asn Leu
 115 120 125

Ile Thr Gly Gly Gln Glu Ala Arg Ala Val Ile Val Lys Ala Phe Gln
 130 135 140

Leu Val Met Gln Lys Cys Ala Ile Phe Glu Pro Lys Ser Ser Ser Phe
 145 150 155 160

Trp Asn Glu Tyr Leu Asn Phe Leu Glu Gln Trp Lys Pro Phe Asn Lys
 165 170 175

Trp Glu Glu Gln Gln Arg Ile Asp Met Leu Arg Glu Phe Tyr Lys Lys
 180 185 190

Met Leu Cys Val Pro Phe Asp Asn Leu Glu Lys Met Trp Asn Arg Tyr
 195 200 205

Thr Gln Trp Glu Gln Glu Ile Asn Ser Leu Thr Ala Arg Lys Phe Ile
 210 215 220

Gly Glu Leu Ser Ala Glu Tyr Met Lys Ala Arg Ser Leu Tyr Gln Glu
 225 230 235 240

Trp Leu Asn Val Thr Asn Gly Leu Lys Arg Ala Ser Pro Ile Asn Leu
 245 250 255

Arg Thr Ala Asn Lys Lys Asn Ile Pro Gln Pro Gly Thr Ser Asp Ser
 260 265 270

Asn Ile Gln Gln Leu Gln Ile Trp Leu Asn Trp Ile Lys Trp Glu Arg
 275 280 285

Glu Asn Lys Leu Met Leu Ser Glu Asp Met Leu Ser Gln Arg Ile Ser
 290 295 300

Tyr Val Tyr Lys Gln Gly Ile Gln Tyr Met Ile Phe Ser Ala Glu Met
 305 310 315 320

Trp Tyr Asp Tyr Ser Met Tyr Ile Ser Glu Asn Ser Asp Arg Gln Asn
 325 330 335

Ile Leu Tyr Thr Ala Leu Leu Ala Asn Pro Asp Ser Pro Ser Leu Thr
 340 345 350

Phe Lys Leu Ser Glu Cys Tyr Glu Leu Asp Asn Asp Ser Glu Ser Val
 355 360 365

Ser Asn Cys Phe Asp Lys Cys Thr Gln Thr Leu Leu Ser Gln Tyr Lys
 370 375 380

Lys Ile Ala Ser Asp Val Asn Ser Gly Glu Asp Asn Asn Thr Glu Tyr
 385 390 395 400

Glu Gln Glu Leu Leu Tyr Lys Gln Arg Glu Lys Leu Thr Phe Val Phe
 405 410 415

Cys Val Tyr Met Asn Thr Met Lys Arg Ile Ser Gly Leu Ser Ala Ala
 420 425 430

Arg Thr Val Phe Gly Lys Cys Arg Lys Leu Lys Arg Ile Leu Thr His
 435 440 445

Asp Val Tyr Val Glu Asn Ala Tyr Leu Glu Phe Gln Asn Gln Asn Asp
 450 455 460

Tyr Lys Thr Ala Phe Lys Val Leu Glu Leu Gly Leu Lys Tyr Phe Gln
 465 470 475 480

Asn Asp Gly Val Tyr Ile Asn Lys Tyr Leu Asp Phe Leu Ile Phe Leu
 485 490 495

Asn Lys Asp Ser Gln Ile Lys Thr Leu Phe Glu Thr Ser Val Glu Lys
 500 505 510

Val Gln Asp Leu Thr Gln Leu Lys Glu Ile Tyr Lys Lys Met Ile Ser
 515 520 525

Tyr Glu Ser Lys Phe Gly Asn Leu Asn Asn Val Tyr Ser Leu Glu Lys
 530 535 540

Arg Phe Phe Glu Arg Phe Pro Gln Glu Asn Leu Ile Glu Val Phe Thr
 545 550 555 560

Ser Arg Tyr Gln Ile Gln Asn Ser Asn Leu Ile Lys Lys Leu Glu Leu
 565 570 575

Thr Tyr Met Tyr Asn Glu Glu Glu Asp Ser Tyr Phe Ser Ser Gly Asn
 580 585 590

Gly Asp Gly His His Gly Ser Tyr Asn Met Ser Ser Ser Asp Arg Lys
 595 600 605

Arg Leu Met Glu Glu Thr Gly Asn Asn Gly Asn Phe Ser Asn Lys Lys
 610 615 620

Phe Lys Arg Asp Ser Glu Leu Pro Thr Glu Val Leu Asp Leu Leu Ser
 625 630 635 640

Val Ile Pro Lys Arg Gln Tyr Phe Asn Thr Asn Leu Leu Asp Ala Gln
 645 650 655

Lys Leu Val Asn Phe Leu Asn Asp Gln Val Glu Ile Pro Thr Val Glu
 660 665 670

Ser Thr Lys Ser Gly
 675

<210> 50

<211> 2000

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 50

atgtccagct ctacgactcc tgatttacta tatccctctg cggacaaagt cgcagagcct	60
agtgacaata tacatggaga tgaactacga cttagagaaa ggattaaaga caatcccacg	120
aatattttat catacttcca gcttattcaa tatttggaag ctcaagagtc atacgctaag	180
gtgagagaag tatacagagca atttcataac acattcccgt tttattcacc tgcgtggact	240
ttgcaactaa aggggtgaatt ggcaagagat gaatttgaga ctggtgagaa gattttggct	300
caatgtcttt ctggcaagtt ggaaaataat gacctatctc tttggtcaac atatttggac	360
tacatacgca gaaaaaaca ctttaattact ggtggacaag aggcgagagc tgttattgtc	420
aaggcattcc aactagttat gcaaaagtgt gcaatttttg aacccaaatc atcttctttt	480
tggaacgaat atctcaattt tttagagcag tggaagccat tcaacaaatg ggaggagcaa	540
cagcgaattg acatgctcag agaattctac aagaaaatgc tatgtgttcc ttttgataat	600
ctagaaaaaa tgtggaatag atacactcaa tgggaacaag aaataaattc cctaacagcc	660
agaaaattta ttggcgagtt atcagccgaa tacatgaaag cccgttcctt ataccaggaa	720
tggttgaacg ttactaatgg attgaaaagg gcatctccaa ttaatctgcg cacagcaaac	780
aagaaaaaca taccacaacc aggtacctca gactcaaaca ttcagcagtt acagatttgg	840
ttgaattgga taaaatggga aaggagagaat aagttgatgc ttagtgaaga tatgctatca	900
caaagaatca gttacgttta taaacaaggt attcaataca tgatatatttc tgctgaaatg	960

tggtagcatt attcaatgta tatatctgaa aattcggatc gacaaaatat cttatatact 1020
 gcgttattag ctaatcccga ctcaccttct cttacattca agttatccga atgctacgaa 1080
 ctggataatg attctgaaag tgtttctaac tgttttgaca agtgactca aactttacta 1140
 tcgcagtata aaaagatcgc ctccgatgta aattcgggtg aagataataa cacagagtat 1200
 gaacaagagc tgctatacaa acagagggaa aaattaacat tcgtgttttg cgtgtatatg 1260
 aatacgatga aaagaatatc aggactatcc gcagcacgta ctgtatttgg taaatgtcgt 1320
 aaactgaagc gtatattaac acatgacgtc tacgtggaaa atgcatattt agaatttcaa 1380
 aatcaaaacg attataagac tgcttttaag gttttagaat tgggtttaaa atacttccaa 1440
 aacgatggag tttatatcaa caaatactta gattttttta tttttttaaa taaggattcg 1500
 cagatcaaaa ccttatttga aacatcagtg gaaaaagtgc aagatttaac ccagctgaag 1560
 gaaatataca agaaaatgat aagttatgaa tcgaaattcg gtaacttaaa caacgtttat 1620
 tctctagaga aaagattttt cgaacgggtc cccaagaaa atttgattga agttttcaca 1680
 agtcgttatc aaattcaaaa ctccaactta ataaagaaat tagagttaac ttatatgtat 1740
 aatgaggaag aagacagtta cttttcttct ggaaacgggg atggccatca tggctcttac 1800
 aatatgagtt cgtcagatag aaagagacta atggaggaaa ctggaaacaa tggaaacttc 1860
 tccaataaga aattcaaaag agactcagag cttccaacag aggttcttga tttattgagc 1920
 gttataccaa aacgtcaata ttttaataca aatttactcg atgcgcagaa attggtgaat 1980
 tttttaaatg atcaagtaga 2000

<210> 51

<211> 296

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 51

Met Asn Arg Gln Ser Gly Val Asn Ala Gly Val Gln Asn Asn Pro Pro
 1 5 10 15

Ser Arg Val Val Tyr Leu Gly Ser Ile Pro Tyr Asp Gln Thr Glu Glu
 20 25 30

Gln Ile Leu Asp Leu Cys Ser Asn Val Gly Pro Val Ile Asn Leu Lys
 35 40 45

Met Met Phe Asp Pro Gln Thr Gly Arg Ser Lys Gly Tyr Ala Phe Ile
 50 55 60

Glu Phe Arg Asp Leu Glu Ser Ser Ala Ser Ala Val Arg Asn Leu Asn
 65 70 75 80

Gly Tyr Gln Leu Gly Ser Arg Phe Leu Lys Cys Gly Tyr Ser Ser Asn
85 90 95

Ser Asp Ile Ser Gly Val Ser Gln Gln Gln Gln Gln Tyr Asn Asn
100 105 110

Ile Asn Gly Asn Asn Asn Asn Asn Gly Asn Asn Asn Asn Asn Ser Asn
115 120 125

Gly Pro Asp Phe Gln Asn Ser Gly Asn Ala Asn Phe Leu Ser Gln Lys
130 135 140

Phe Pro Glu Leu Pro Ser Gly Ile Asp Val Asn Ile Asn Met Thr Thr
145 150 155 160

Pro Ala Met Met Ile Ser Ser Glu Leu Ala Lys Lys Pro Lys Glu Val
165 170 175

Gln Leu Lys Phe Leu Gln Lys Phe Gln Glu Trp Thr Arg Ala His Pro
180 185 190

Glu Asp Ala Val Ser Leu Leu Glu Leu Cys Pro Gln Leu Ser Phe Val
195 200 205

Thr Ala Glu Leu Leu Leu Thr Asn Gly Ile Cys Lys Val Asp Asp Leu
210 215 220

Ile Pro Leu Ala Ser Arg Pro Gln Glu Glu Ala Ser Ala Thr Asn Asn
225 230 235 240

Asn Ser Val Asn Glu Val Val Asp Pro Ala Val Leu Asn Lys Gln Lys
245 250 255

Glu Leu Leu Lys Gln Val Leu Gln Leu Asn Asp Ser Gln Ile Ser Ile
260 265 270

Leu Pro Asp Asp Glu Arg Met Ala Ile Trp Asp Leu Lys Gln Lys Ala
275 280 285

Leu Arg Gly Glu Phe Gly Ala Phe
290 295

<210> 52

<211> 891

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 52
 atgaataggc agagcgggtgt gaatgcgggg gtacagaaca acccaccatc ccgagtgggtg 60
 tatctgggtt ctataccata cgatcaaaca gaggagcaga tacttgattt atgtagtaat 120
 gttgggcccg tgatcaattt gaaaatgatg ttcgaccccc aaactggtag gtcgaaaggg 180
 tacgcgttta ttgaatttag agatttagag tccagtgcc a gcgcagtacg taatttgaat 240
 ggataccaat taggctctag gtttttgaaa tgcggttact ccagcaatag tgatatatcg 300
 ggagtttcac aacagcaaca acaacagtac aacaacatta atgggaacaa taacaacaat 360
 ggaaataata ataataatag taatgggccg gactttcaaa acagcggaaa tgccaatttt 420
 ctaagtcaaa agtttccaga attgccctct ggtatcgacg ttaacataaa catgaccacc 480
 cctgctatga tgatatcgag cgaactagct aaaaaaccga aagaggtgca gttgaaattt 540
 ttacaaaaat tccaagaatg gacaagagcg catcctgaag atgctgtttc gctattagag 600
 ctgtgtccac agttgagttt tgttacggct gaattattgc taacgaatgg gatatgtaaa 660
 gtggatgatt tgatcccggt agcttccagg ccgcaagaag aggcacggc tacgaataac 720
 aatagcgtga acgaggtggg ggatccagct gtgcttaaca aacagaaaga actactgaaa 780
 caggtgttac aactgaatga cagtcaaatt tctatcttgc ccgatgatga aaggatggct 840
 atttgggact taaaacaaaa agcattaagg ggagaatttg gtgcattttg a 891

<210> 53
 <211> 297
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 53

Met Val Val Ile Ala Asn Ala His Asn Glu Leu Ile His Asp Ala Val
 1 5 10 15

Leu Asp Tyr Tyr Gly Lys Arg Leu Ala Thr Cys Ser Ser Asp Lys Thr
 20 25 30

Ile Lys Ile Phe Glu Val Glu Gly Glu Thr His Lys Leu Ile Asp Thr
 35 40 45

Leu Thr Gly His Glu Gly Pro Val Trp Arg Val Asp Trp Ala His Pro
 50 55 60

Lys Phe Gly Thr Ile Leu Ala Ser Cys Ser Tyr Asp Gly Lys Val Leu
 65 70 75 80

Ile Trp Lys Glu Glu Asn Gly Arg Trp Ser Gln Ile Ala Val His Ala
 85 90 95

87/148

Val His Ser Ala Ser Val Asn Ser Val Gln Trp Ala Pro His Glu Tyr
 100 105 110

Gly Pro Leu Leu Leu Val Ala Ser Ser Asp Gly Lys Val Ser Val Val
 115 120 125

Glu Phe Lys Glu Asn Gly Thr Thr Ser Pro Ile Ile Ile Asp Ala His
 130 135 140

Ala Ile Gly Val Asn Ser Ala Ser Trp Ala Pro Ala Thr Ile Glu Glu
 145 150 155 160

Asp Gly Glu His Asn Gly Thr Lys Glu Ser Arg Lys Phe Val Thr Gly
 165 170 175

Gly Ala Asp Asn Leu Val Lys Ile Trp Lys Tyr Asn Ser Asp Ala Gln
 180 185 190

Thr Tyr Val Leu Glu Ser Thr Leu Glu Gly His Ser Asp Trp Val Arg
 195 200 205

Asp Val Ala Trp Ser Pro Thr Val Leu Leu Arg Ser Tyr Leu Ala Ser
 210 215 220

Val Ser Gln Asp Arg Thr Cys Ile Ile Trp Thr Gln Asp Asn Glu Gln
 225 230 235 240

Gly Pro Trp Lys Lys Thr Leu Leu Lys Glu Glu Lys Phe Pro Asp Val
 245 250 255

Leu Trp Arg Ala Ser Trp Ser Leu Ser Gly Asn Val Leu Ala Leu Ser
 260 265 270

Gly Gly Asp Asn Lys Val Thr Leu Trp Lys Glu Asn Leu Glu Gly Lys
 275 280 285

Trp Glu Pro Ala Gly Glu Val His Gln
 290 295

<210> 54

<211> 894

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 54

atggtcgtca tagctaatagc gcacaacgaa ttaatccatg acgctgttct agactattat 60

gggaagcgcc ttgcaacctg ctcttctgac aagacaatca agatctttga agtcgaagga 120

gaaacacaca agttaataga cacgttgact ggccacgaag gccagtttg gcgtgttgat 180

tgggcacatc ctaaattcgg aaccattttg gcatcgtgtt cttatgatgg taaagtgttg 240
 atttggaagg aagaaaacgg tagatggtct caaattgccg ttcattgctgt ccactctgct 300
 tctgtcaact ctgttcaatg ggctcctcat gaatatggcc ccctactgct gggtgcttcc 360
 tctgatggta aggtctccgt agtagagtgc aaagaaaacg gtactacttc cccaataatc 420
 atcgatgctc atgccattgg cgtaaactct gcttcttggg ctccagctac catcgaagaa 480
 gatggtgaac acaacgggtac taaagaatct cgcaagtttg ttactggggg tgctgacaat 540
 ttggtaaaga tttggaagta caattcagat gcccaaactt atgttctgga aagcacctta 600
 gaaggtcaca gcgattgggt tagagacgta gcatgggtcac ctactgttct tctacgttct 660
 tatttggcca gtgtttctca agatcgacc tgtattattt ggactcaaga caatgaacaa 720
 ggcccatgga aaaaaacttt attaaaagaa gaaaaattcc cagatgtttt atggagagcc 780
 agttggtctt tgcaggttaa tgtactagct ctttccggtg gcgataataa agttacttta 840
 tggaaggaaa atcttgaggg taaatgggaa cccgctggtg aagttcatca gtga 894

<210> 55

<211> 1273

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 55

Met Val Lys Leu Ala Glu Phe Ser Arg Thr Ala Thr Phe Ala Trp Ser
 1 5 10 15

His Asp Lys Ile Pro Leu Leu Val Ser Gly Thr Val Ser Gly Thr Val
 20 25 30

Asp Ala Asn Phe Ser Thr Asp Ser Ser Leu Glu Leu Trp Ser Leu Leu
 35 40 45

Ala Ala Asp Ser Glu Lys Pro Ile Ala Ser Leu Gln Val Asp Ser Lys
 50 55 60

Phe Asn Asp Leu Asp Trp Ser His Asn Asn Lys Ile Ile Ala Gly Ala
 65 70 75 80

Leu Asp Asn Gly Ser Leu Glu Leu Tyr Ser Thr Asn Glu Ala Asn Asn
 85 90 95

Ala Ile Asn Ser Met Ala Arg Phe Ser Asn His Ser Ser Ser Val Lys
 100 105 110

Thr Val Lys Phe Asn Ala Lys Gln Asp Asn Val Leu Ala Ser Gly Gly
 115 120 125

Asn Asn Gly Glu Ile Phe Ile Trp Asp Met Asn Lys Cys Thr Glu Ser
 130 135 140

Pro Ser Asn Tyr Thr Pro Leu Thr Pro Gly Gln Ser Met Ser Ser Val
 145 150 155 160

Asp Glu Val Ile Ser Leu Ala Trp Asn Gln Ser Leu Ala His Val Phe
 165 170 175

Ala Ser Ala Gly Ser Ser Asn Phe Ala Ser Ile Trp Asp Leu Lys Ala
 180 185 190

Lys Lys Glu Val Ile His Leu Ser Tyr Thr Ser Pro Asn Ser Gly Ile
 195 200 205

Lys Gln Gln Leu Ser Val Val Glu Trp His Pro Lys Asn Ser Thr Arg
 210 215 220

Val Ala Thr Ala Thr Gly Ser Asp Asn Asp Pro Ser Ile Leu Ile Trp
 225 230 235 240

Asp Leu Arg Asn Ala Asn Thr Pro Leu Gln Thr Leu Asn Gln Gly His
 245 250 255

Gln Lys Gly Ile Leu Ser Leu Asp Trp Cys His Gln Asp Glu His Leu
 260 265 270

Leu Leu Ser Ser Gly Arg Asp Asn Thr Val Leu Leu Trp Asn Pro Glu
 275 280 285

Ser Ala Glu Gln Leu Ser Gln Phe Pro Ala Arg Gly Asn Trp Cys Phe
 290 295 300

Lys Thr Lys Phe Ala Pro Glu Ala Pro Asp Leu Phe Ala Cys Ala Ser
 305 310 315 320

Phe Asp Asn Lys Ile Glu Val Gln Thr Leu Gln Asn Leu Thr Asn Thr
 325 330 335

Leu Asp Glu Gln Glu Thr Glu Thr Lys Gln Gln Glu Ser Glu Thr Asp
 340 345 350

Phe Trp Asn Asn Val Ser Arg Glu Glu Ser Lys Glu Lys Pro Ser Val
 355 360 365

Phe His Leu Gln Ala Pro Thr Trp Tyr Gly Glu Pro Ser Pro Ala Ala

370		375		380
His Trp Ala Phe Gly Gly Lys Leu Val Gln Ile Thr Pro Asp Gly Lys				
385		390		395 400
Gly Val Ser Ile Thr Asn Pro Lys Ile Ser Gly Leu Glu Ser Asn Thr				
	405		410	415
Thr Leu Ser Glu Ala Leu Lys Thr Lys Asp Phe Lys Pro Leu Ile Asn				
	420		425	430
Gln Arg Leu Val Lys Val Ile Asp Asp Val Asn Glu Glu Asp Trp Asn				
	435		440	445
Leu Leu Glu Lys Leu Ser Met Asp Gly Thr Glu Glu Phe Leu Lys Glu				
	450		455	460
Ala Leu Ala Phe Asp Asn Asp Glu Ser Asp Ala Gln Asp Asp Ala Asn				
465		470		475 480
Asn Glu Lys Glu Asp Asp Gly Glu Glu Phe Phe Gln Gln Ile Glu Thr				
	485		490	495
Asn Phe Gln Pro Glu Gly Asp Phe Ser Leu Ser Gly Asn Ile Glu Gln				
	500		505	510
Thr Ile Ser Lys Asn Leu Val Ser Gly Asn Ile Lys Ser Ala Val Lys				
	515		520	525
Asn Ser Leu Glu Asn Asp Leu Leu Met Glu Ala Met Val Ile Ala Leu				
	530		535	540
Asp Ser Asn Asn Glu Arg Leu Lys Glu Ser Val Lys Asn Ala Tyr Phe				
545		550		555 560
Ala Lys Tyr Gly Ser Lys Ser Ser Leu Ser Arg Ile Leu Tyr Ser Ile				
	565		570	575
Ser Lys Arg Glu Val Asp Asp Leu Val Glu Asn Leu Asp Val Ser Gln				
	580		585	590
Trp Lys Phe Ile Ser Lys Ala Ile Gln Asn Leu Tyr Pro Asn Asp Ile				
	595		600	605
Ala Gln Arg Asn Glu Met Leu Ile Lys Leu Gly Asp Arg Leu Lys Glu				
	610		615	620

Ile Gln Pro Gly Met Pro Ile Met Gly Asp Tyr Asn Ala Gln Ser Ser
885 890 895

Ser Ile Pro Ser Gln Pro Pro Ile Asn Ala Val Ser Gly Gln Thr Pro
900 905 910

His Leu Asn Arg Lys Ala Asn Asp Gly Trp Asn Asp Leu Pro Leu Lys
915 920 925

Val Lys Glu Lys Pro Ser Arg Ala Lys Ala Val Ser Val Ala Pro Pro
930 935 940

Asn Ile Leu Ser Thr Pro Thr Pro Leu Asn Gly Ile Pro Ala Asn Ala
945 950 955 960

Ala Ser Thr Met Pro Pro Pro Pro Leu Ser Arg Ala Pro Ser Ser Val
965 970 975

Ser Met Val Ser Pro Pro Pro Leu His Lys Asn Ser Arg Val Pro Ser
980 985 990

Leu Val Ala Thr Ser Glu Ser Pro Arg Ala Ser Ile Ser Asn Pro Tyr
995 1000 1005

Ala Pro Pro Gln Ser Ser Gln Gln Phe Pro Ile Gly Thr Ile Ser
1010 1015 1020

Thr Ala Asn Gln Thr Ser Asn Thr Ala Gln Val Ala Ser Ser Asn
1025 1030 1035

Pro Tyr Ala Pro Pro Pro Gln Gln Arg Val Ala Thr Pro Leu Ser
1040 1045 1050

Gly Gly Val Pro Pro Ala Pro Leu Pro Lys Ala Ser Asn Pro Tyr
1055 1060 1065

Ala Pro Thr Ala Thr Thr Gln Pro Asn Gly Ser Ser Tyr Pro Pro
1070 1075 1080

Thr Gly Pro Tyr Thr Asn Asn His Thr Met Thr Ser Pro Pro Pro
1085 1090 1095

Val Phe Asn Lys Pro Pro Thr Gly Pro Pro Pro Ile Ser Met Lys
1100 1105 1110

Lys Arg Ser Asn Lys Leu Ala Ser Ile Glu Gln Asn Pro Ser Gln
1115 1120 1125

Gly Ala Thr Tyr Pro Pro Thr Leu Ser Ser Ser Ala Ser Pro Leu
1130 1135 1140

Gln Pro Ser Gln Pro Pro Thr Leu Ala Ser Gln Val Asn Thr Ser
1145 1150 1155

Ala Glu Asn Val Ser His Glu Ile Pro Ala Asp Gln Gln Pro Ile
1160 1165 1170

Val Asp Phe Leu Lys Glu Glu Leu Ala Arg Val Thr Pro Leu Thr
1175 1180 1185

Pro Lys Glu Tyr Ser Lys Gln Leu Lys Asp Cys Asp Lys Arg Leu
1190 1195 1200

Lys Ile Leu Phe Tyr His Leu Glu Lys Gln Asp Leu Leu Thr Gln
1205 1210 1215

Pro Thr Ile Asp Cys Leu His Asp Leu Val Ala Leu Met Lys Glu
1220 1225 1230

Lys Lys Tyr Lys Glu Ala Met Val Ile His Ala Asn Ile Ala Thr
1235 1240 1245

Asn His Ala Gln Glu Gly Gly Asn Trp Leu Thr Gly Val Lys Arg
1250 1255 1260

Leu Ile Gly Ile Ala Glu Ala Thr Leu Asn
1265 1270

<210> 56

<211> 2000

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 56

```

atggtcaaac ttgctgagtt ttctcgaaca gccacgtttg cgtggtcaca tgataaaatt      60
ccattattgg tctctggtac cgtatctggt acggtggatg ctaatttctc cactgattca      120
tctctagaat tgtggtcatt gttggctgct gattcggaga agcctattgc ttccttgcaa      180
gtggattcca aattcaatga tttggattgg tctcataata acaagattat tgctgggtgct      240
ctggataacg gtagtttggg attgtactcc accaatgaag caaacaacgc tatcaactcc      300
atggccagat ttagcaacca ttcttcctct gtgaagacgg taaagtttaa cgcaaagcaa      360
gacaacgttc ttgcttcggg tggtacaac ggtgaaattt ttatttgga catgaataaa      420
tgcactgaat cgccctccaa ttatactcca ttgacaccgg gtcaatcgat gtcgtccggt      480

```

gacgaggtca tttccctagc atggaaccaa tctttggccc atgtttttgc atctgccggg 540
 tcgtctaatt tcgcatctat ttgggatttg aaggctaaga aggaagtcac tcatctaagt 600
 tacacttcac ctaattcagg tatcaagcaa cagctgtccg ttgttgaatg gcacccaaaa 660
 aactccacaa gagtggcaac ggctactggt agcgataatg atccatctat cctgatctgg 720
 gatttaagaa acgccaacac accattgcag actttaaatc aaggccatca aaagggtatt 780
 ttgtcattag attggtgtca tcaggacgaa catctattat tgtccagtgg tagagataat 840
 accgttcttc tatggaaccc tgagtcagcc gaacaactgt cccaattccc agctcgtgga 900
 aactggtggt ttaagaccaa atttgcacca gaggtccag acctatttgc ttgtgcctcc 960
 tttgataaca aaattgaggt acagactttg caaaatctca caaacacttt ggatgagcaa 1020
 gaaaccgaaa ctaagcagca agaactctgaa acagattttt ggaataatgt ttcccagagag 1080
 gaatcaaaag agaagccatc tgttttccat ttacaagccc caacttggtg tggggaacca 1140
 tctcccgag ctcattgggc tttcgggtgt aaattgggtc aaattactcc agatggtaaa 1200
 ggtgtatcta taacaaaccc aaaaatttca ggcttagaat caaacactac tttgagtga 1260
 gcgttgaaaa ctaaggattt caaacatta ataatcaaa gactgggtcaa agttattgat 1320
 gacgttaatg aagaagattg gaatttattg gaaaagttat caatggacgg tactgaggag 1380
 ttcttgaaag aggctcttgc attcgacaac gatgaatcag atgcacaaga cgatgccaac 1440
 aatgagaaag aagacgatgg ggaagaattc tttcaacaaa ttgaaaccaa tttccaaccc 1500
 gagggcgatt tctccttgtc tggtaatatc gaacaaacta tttccaagaa cttgggttct 1560
 ggcaacatta agagcgtgt gaaaaattct ctagagaatg acttactaat ggaggccatg 1620
 gtgatcgcat tagattcaaa taacgaaaga ttaaaggaaa gtgtcaagaa tgcctatatt 1680
 gcgaagtatg gatctaaatc atcgctctcg aggatactat actccatttc taagagggaa 1740
 gtagatgatt tgggtgaaaa tttggatgtc tctcagtggg agtttatctc taaagcaatt 1800
 caaaacttat atccaaatga tatcgcccag aggaatgaaa tgttgattaa attgggagac 1860
 aggttaaagg aaaatgggtc tagacaagat tctttgactt tgtacttggc tgccggatca 1920
 ttagataagg tggcttcaat ttggttatca gaatttccag atttgaggga taaattgaag 1980
 aaagataata agacaattta 2000

<210> 57

<211> 649

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 57

Met Ser Arg Ala Val Gly Ile Asp Leu Gly Thr Thr Tyr Ser Cys Val
 1 5 10 15

Ala His Phe Ser Asn Asp Arg Val Glu Ile Ile Ala Asn Asp Gln Gly
 20 25 30

Asn Arg Thr Thr Pro Ser Tyr Val Ala Phe Thr Asp Thr Glu Arg Leu
 35 40 45

Ile Gly Asp Ala Ala Lys Asn Gln Ala Ala Ile Asn Pro His Asn Thr
 50 55 60

Val Phe Asp Ala Lys Arg Leu Ile Gly Arg Lys Phe Asp Asp Pro Glu
 65 70 75 80

Val Thr Thr Asp Ala Lys His Phe Pro Phe Lys Val Ile Ser Arg Asp
 85 90 95

Gly Lys Pro Val Val Gln Val Glu Tyr Lys Gly Glu Thr Lys Thr Phe
 100 105 110

Thr Pro Glu Glu Ile Ser Ser Met Val Leu Ser Lys Met Lys Glu Thr
 115 120 125

Ala Glu Asn Tyr Leu Gly Thr Thr Val Asn Asp Ala Val Val Thr Val
 130 135 140

Pro Ala Tyr Phe Asn Asp Ser Gln Arg Gln Ala Thr Lys Asp Ala Gly
 145 150 155 160

Thr Ile Ala Gly Met Asn Val Leu Arg Ile Ile Asn Glu Pro Thr Ala
 165 170 175

Ala Ala Ile Ala Tyr Gly Leu Asp Lys Lys Gly Arg Ala Glu His Asn
 180 185 190

Val Leu Ile Phe Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Leu Leu
 195 200 205

Ser Ile Asp Glu Gly Val Phe Glu Val Lys Ala Thr Ala Gly Asp Thr
 210 215 220

His Leu Gly Gly Glu Asp Phe Asp Asn Arg Leu Val Asn His Leu Ala
 225 230 235 240

Thr Glu Phe Lys Arg Lys Thr Lys Lys Asp Ile Ser Asn Asn Gln Arg
 245 250 255

Ser Leu Arg Arg Leu Arg Thr Ala Ala Glu Arg Ala Lys Arg Ala Leu

260	265	270
Ser Ser Ser Ser Gln Thr Ser Ile Glu Ile Asp Ser Leu Phe Glu Gly		
275	280	285
Met Asp Phe Tyr Thr Ser Leu Thr Arg Ala Arg Phe Glu Glu Leu Cys		
290	295	300
Ala Asp Leu Phe Arg Ser Thr Leu Glu Pro Val Glu Lys Val Leu Lys		
305	310	315
Asp Ser Lys Leu Asp Lys Ser Gln Ile Asp Glu Ile Val Leu Val Gly		
325	330	335
Gly Ser Thr Arg Ile Pro Lys Ile Gln Lys Leu Val Ser Asp Phe Phe		
340	345	350
Asn Gly Lys Glu Pro Asn Arg Ser Ile Asn Pro Asp Glu Ala Val Ala		
355	360	365
Tyr Gly Ala Ala Val Gln Ala Ala Ile Leu Thr Gly Asp Gln Ser Thr		
370	375	380
Lys Thr Gln Asp Leu Leu Leu Leu Asp Val Ala Pro Leu Ser Leu Gly		
385	390	395
Ile Glu Thr Ala Gly Gly Ile Met Thr Lys Leu Ile Pro Arg Asn Ser		
405	410	415
Thr Ile Pro Thr Lys Lys Ser Glu Thr Phe Ser Thr Tyr Ala Asp Asn		
420	425	430
Gln Pro Gly Val Leu Ile Gln Val Phe Glu Gly Glu Arg Thr Arg Thr		
435	440	445
Lys Asp Asn Asn Leu Leu Gly Lys Phe Glu Leu Ser Gly Ile Pro Pro		
450	455	460
Ala Pro Arg Gly Val Pro Gln Ile Asp Val Thr Phe Asp Ile Asp Ala		
465	470	475
Asn Gly Ile Leu Asn Val Ser Ala Leu Glu Lys Gly Thr Gly Lys Ser		
485	490	495
Asn Lys Ile Thr Ile Thr Asn Asp Lys Gly Arg Leu Ser Lys Asp Asp		
500	505	510

97/148

Ile Asp Arg Met Val Ser Glu Ala Glu Lys Tyr Arg Ala Asp Asp Glu
515 520 525

Arg Glu Ala Glu Arg Val Gln Ala Lys Asn Gln Leu Glu Ser Tyr Ala
530 535 540

Phe Thr Leu Lys Asn Thr Ile Asn Glu Ala Ser Phe Lys Glu Lys Val
545 550 555 560

Gly Glu Asp Asp Ala Lys Arg Leu Glu Thr Ala Ser Gln Glu Thr Ile
565 570 575

Asp Trp Leu Asp Ala Ser Gln Ala Ala Ser Thr Asp Glu Tyr Lys Asp
580 585 590

Arg Gln Lys Glu Leu Glu Gly Ile Ala Asn Pro Ile Met Thr Lys Phe
595 600 605

Tyr Gly Ala Gly Ala Gly Ala Gly Pro Gly Ala Gly Glu Ser Gly Gly
610 615 620

Phe Pro Gly Ser Met Pro Asn Ser Gly Ala Thr Gly Gly Gly Glu Asp
625 630 635 640

Thr Gly Pro Thr Val Glu Glu Val Asp
645

<210> 58

<211> 1950

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 58

```

atgtctagag cagttggtat tgatttggga acaacttact cgtgtgttgc tcatttttcc      60
aatgataggg tagagataat tgcaaatgat caaggtaata ggaccactcc atcgtatgtg      120
gctttcacag acaccgaaag attaattggt gacgccgcca aaaatcaagc tgcaatcaat      180
cctcataata cagtttttga tgcaaagcgg ttaattggtc gtaaatttga tgatcctgaa      240
gtgacgacag atgccaagca cttccctttc aaagttatat ccagagatgg taaacctgta      300
gtgcaagtag aatataaggg tgaaacgaaa acatttacgc ctgaggaaat ttcttccatg      360
gttttaagca aaatgaagga aactgctgag aactatttgg gaactacggt caatgatgct      420
gttgtaactg ttcttgcata tttcaatgat tctcaaagac aagccactaa ggatgcagga      480
actattgcag ggatgaacgt ttacgtatt atcaatgaac ccactgcagc agcaattgct      540
tatggccttg ataagaaagg cagggctgag cacaatgtcc tgatttttga tttgggtggg      600
ggttactttt acgtctcttt actttcaatt gatgaggggt tttttgaggt taaggctacc      660

```

gcaggagaca ctcatttagg tggatgaagat ttgataata ggttggtgaa ccatttagcc 720
 actgaattca aaaggaaaac gaaaaaggac atctctaata atcaaagatc gttaagaaga 780
 ttgagaactg cggcagaaaag agctaagaga gcgctttctt cctcatctca aacctcgatc 840
 gagatcgatt ctttatttga aggtatggat ttctacactt cgtaacaag ggcaagggtt 900
 gaagagctat gtgctgattt attcagatcc acattggaac cagtagaaaa ggttccttaa 960
 gattcgaagc tggacaagtc ccaaattgat gagattgtgt tagtcgggtg atctaccaga 1020
 atcccaaaga ttcagaaatt agtttctgac ttcttcaatg gcaaagagcc taatcgttct 1080
 atcaaccggg atgaggctgt tgcttatggg gcagcgttc aagctgccat ttaaccggc 1140
 gatcaatcaa caaagacaca agatttacta ttattggatg ttgcgccatt gtccctagga 1200
 attgaaactg caggcggcat aatgactaag ctaattccta gaaactcaac gattccaaca 1260
 aagaaatcgg aaaccttctc tacctatgca gataatcaac ctggtgtttt aattcaagtc 1320
 tttgaagggtg aaagaacaag acaaaaggat aataacttac ttggtaaatt cgaattaagt 1380
 ggcattccgc ctgctcccag aggtgtgcct caaattgatg ttaccttga tatcgacgct 1440
 aatggtattc ttaatgtgtc tgctttggaa aagggtactg gtaagagtaa caaatcacg 1500
 atcactaacg ataaaggtag gctctcgaag gatgatattg ataggatggg ttctgaagct 1560
 gaaaaatata gggctgacga tgaaaggag gcagaacgag ttcaggctaa gaatcagctt 1620
 gaatcgtatg catttacttt gaagaatacc ataaacgaag caagtttcaa agagaaagta 1680
 ggtgaagatg atgcaaagag attagaaaca gcgtctcagg aaaccattga ctggttagat 1740
 gcacgcagg cagcctctac ggacgaatat aaggatagac aaaaggagt ggaaggcatt 1800
 gccaatccaa taatgacgaa attttacggg gctggtgccg gcgcagggtc tggagcggg 1860
 gaatccggtg gattccccgg atccatgcc aactcgggtg ctacgggagg tggagaagat 1920
 acagggtcaa cagtgaaga ggttgattga 1950

<210> 59

<211> 206

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 59

Met Pro Ser His Arg Asn Ser Asn Leu Lys Phe Cys Thr Val Cys Ala
 1 5 10 15

Ser Asn Asn Asn Arg Ser Met Glu Ser His Lys Val Leu Gln Glu Ala
 20 25 30

Gly Tyr Asn Val Ser Ser Tyr Gly Thr Gly Ser Ala Val Arg Leu Pro
 35 40 45

Gly Leu Ser Ile Asp Lys Pro Asn Val Tyr Ser Phe Gly Thr Pro Tyr
50 55 60

Asn Asp Ile Tyr Asn Asp Leu Leu Ser Gln Ser Ala Asp Arg Tyr Lys
65 70 75 80

Ser Asn Gly Leu Leu Gln Met Leu Asp Arg Asn Arg Arg Leu Lys Lys
85 90 95

Ala Pro Glu Lys Trp Gln Glu Ser Thr Lys Val Phe Asp Phe Val Phe
100 105 110

Thr Cys Glu Glu Arg Cys Phe Asp Ala Val Cys Glu Asp Leu Met Asn
115 120 125

Arg Gly Gly Lys Leu Asn Lys Ile Val His Val Ile Asn Val Asp Ile
130 135 140

Lys Asp Asp Asp Glu Asn Ala Lys Ile Gly Ser Lys Ala Ile Leu Glu
145 150 155 160

Leu Ala Asp Met Leu Asn Asp Lys Ile Glu Gln Cys Glu Lys Asp Asp
165 170 175

Ile Pro Phe Glu Asp Cys Ile Met Asp Ile Leu Thr Glu Trp Gln Ser
180 185 190

Ser His Ser Gln Leu Pro Ser Leu Tyr Ala Pro Ser Tyr Tyr
195 200 205

<210> 60

<211> 621

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 60

atgcctagtc atcgcaattc aaacttgaag ttttgcacag tttgtgcatc aaacaacaat 60

cgttcaatgg aatcgcataa agtcctgcaa gaagcaggct ataatgtag ctcttacgga 120

acagggttcag ctgtgagact gcctgggtcta tcgatagata agcctaattgt gtactcattt 180

ggtagaccct ataatgatat atataatgat cttttatcac aatcagcaga ccgttacaag 240

tcgaacgggt tattgcaa at gctggatcgt aatagaagac tcaaaaaagc acctgaaaaa 300

tggcaagaaa gtacaaaagt cttcgacttc gttttcactt gtgaagagag atgttttgat 360

gccgtttgtg aagatttgat gaatagagggt gggaaattaa acaaaatagt gcatgtaatt 420

aatgttgaca ttaaagatga tgatgaaaat gctaaaattg gtagcaaagc tatattggaa 480

ttagctgata tgctcaatga taaaatagaa caatgtgaaa aagatgacat tccctttgaa 540
 gattgtataa tggacatttt aactgagtgg caaagctcac attctcaact accgtcatta 600
 tacgctcctt catattacta a 621

<210> 61
 <211> 516
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 61

Met Ser Thr Gln Gln Gln Ser Tyr Thr Ile Trp Ser Pro Gln Asp Thr
1 5 10 15

Val Lys Asp Val Ala Glu Ser Leu Gly Leu Glu Asn Ile Asn Asp Asp
20 25 30

Val Leu Lys Ala Leu Ala Met Asp Val Glu Tyr Arg Ile Leu Glu Ile
35 40 45

Ile Glu Gln Ala Val Lys Phe Lys Arg His Ser Lys Arg Asp Val Leu
50 55 60

Thr Thr Asp Asp Val Ser Lys Ala Leu Arg Val Leu Asn Val Glu Pro
65 70 75 80

Leu Tyr Gly Tyr Tyr Asp Gly Ser Glu Val Asn Lys Ala Val Ser Phe
85 90 95

Ser Lys Val Asn Thr Ser Gly Gly Gln Ser Val Tyr Tyr Leu Asp Glu
100 105 110

Glu Glu Val Asp Phe Asp Arg Leu Ile Asn Glu Pro Leu Pro Gln Val
115 120 125

Pro Arg Leu Pro Thr Phe Thr Thr His Trp Leu Ala Val Glu Gly Val
130 135 140

Gln Pro Ala Ile Ile Gln Asn Pro Asn Leu Asn Asp Ile Arg Val Ser
145 150 155 160

Gln Pro Pro Phe Ile Arg Gly Ala Ile Val Thr Ala Leu Asn Asp Asn
165 170 175

Ser Leu Gln Thr Pro Val Thr Ser Thr Thr Ala Ser Ala Ser Val Thr
180 185 190

Asp Thr Gly Ala Ser Gln His Leu Ser Asn Val Lys Pro Gly Gln Asn
 195 200 205

Thr Glu Val Lys Pro Leu Val Lys His Val Leu Ser Lys Glu Leu Gln
 210 215 220

Ile Tyr Phe Asn Lys Val Ile Ser Thr Leu Thr Ala Lys Ser Gln Ala
 225 230 235 240

Asp Glu Ala Ala Gln His Met Lys Gln Ala Ala Leu Thr Ser Leu Arg
 245 250 255

Thr Asp Ser Gly Leu His Gln Leu Val Pro Tyr Phe Ile Gln Phe Ile
 260 265 270

Ala Glu Gln Ile Thr Gln Asn Leu Ser Asp Leu Gln Leu Leu Thr Thr
 275 280 285

Ile Leu Glu Met Ile Tyr Ser Leu Leu Ser Asn Thr Ser Ile Phe Leu
 290 295 300

Asp Pro Tyr Ile His Ser Leu Met Pro Ser Ile Leu Thr Leu Leu Leu
 305 310 315 320

Ala Lys Lys Leu Gly Gly Ser Pro Lys Asp Asp Ser Pro Gln Glu Ile
 325 330 335

His Glu Phe Leu Glu Arg Thr Asn Ala Leu Arg Asp Phe Ala Ala Ser
 340 345 350

Leu Leu Asp Tyr Val Leu Lys Lys Phe Pro Gln Ala Tyr Lys Ser Leu
 355 360 365

Lys Pro Arg Val Thr Arg Thr Leu Leu Lys Thr Phe Leu Asp Ile Asn
 370 375 380

Arg Val Phe Gly Thr Tyr Tyr Gly Cys Leu Lys Gly Val Ser Val Leu
 385 390 395 400

Glu Gly Glu Ser Ile Arg Phe Phe Leu Gly Asn Leu Asn Asn Trp Ala
 405 410 415

Arg Leu Val Phe Asn Glu Ser Gly Ile Thr Leu Asp Asn Ile Glu Glu
 420 425 430

His Leu Asn Asp Asp Ser Asn Pro Thr Arg Thr Lys Phe Thr Lys Glu
 435 440 445

Glu Thr Gln Ile Leu Val Asp Thr Val Ile Ser Ala Leu Leu Val Leu
450 455 460

Lys Lys Asp Leu Pro Asp Leu Tyr Glu Gly Lys Gly Glu Lys Val Thr
465 470 475 480

Asp Glu Asp Lys Glu Lys Leu Leu Glu Arg Cys Gly Val Thr Ile Gly
485 490 495

Phe His Ile Leu Lys Arg Asp Asp Ala Lys Glu Leu Ile Ser Ala Ile
500 505 510

Phe Phe Gly Glu
515

<210> 62
<211> 1551
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 62
atgtctacac agcagcaatc ttacacaatt tgggtctccac aagatactgt taaagatgtc 60
gctgagtcgc ttgggttgga aaatatcaac gatgacgtgt tgaaagcact tgctatggac 120
gttgaatacc gtattctaga gattattgaa caggcagtca agtttaaaag acactctaaa 180
agagatgttt taactacaga cgatgtatcg aaggccctgc gtgttctgaa tgtggaacct 240
ttatatggat attatgatgg ctctgaagtt aataaagctg tatcatttag taaggttaac 300
acaagtggag gacagtcggt ttactacctg gatgaggaag aagtagattt tgatagatta 360
ataaatgagc cattaccgca agtgcctcgt ctaccaactt ttactacaca ttggctagca 420
gttgaagggg ttcaacctgc cataatccaa aatccaaatt tgaatgatat aagggtatcc 480
caaccacat ttattagagg tgctattgtc actgctttga atgataacag cctccaaaca 540
cctgtaacgt cgacgacagc aagtgccttct gtgacggata caggcgcctc tcaacacctg 600
tctaattgta aaccaggaca gaatactgaa gtcaaacctt tagtgaaaca cgttttatcc 660
aaagaattac agatttatth caataaggct atttcaacct tgacagcgaa gagtcaagct 720
gatgaagctg cacagcatat gaaacaagca gctttgactt cactacgaac tgatagtggc 780
ctgcaccaac tggttccata ttttattcaa tttattgccg aacaaatcac acaaaatctt 840
tccgatttac aattgctaac aacaattttg gaaatgattt actctttact aagcaatact 900
tctattttcc tggaccctta tttcattcg ttaatgcctt ccattttaac cttgctatta 960
gcaaaaaaac ttggtggatc accaaaagat gattcgccgc aagaaattca tgaattttta 1020
gaaagaacga atgcactacg tgattttgcc gcacttttgt tggactatgt attgaaaaaa 1080


```

tttcctcaag catacaaatc cctgaagcca agagtaacta ggacactgct aaagacattt 1140
ttggacataa accgtgtttt tgggacttac tatgggtgct taaaggggtgt atcggtacta 1200
gaaggtgaat ccatcagatt cttcttagga aacctaaata attgggcgcg cttggtattc 1260
aatgaaagcg gaataactct agacaatata gaggaacatt tgaacgatga ctccaatcca 1320
actaggacca aatttactaa agaagaaact caaatcttgg ttgatactgt aattagtgca 1380
ttgttagttc taaaaaagga ttaccagat ttgtacgagg gtaaagggtga aaaagttacg 1440
gatgaggata aggaaaaaatt gctagagagg tgtggtgtca ctattggatt tcacattttg 1500
aaaagagacg acgcaaaaga attaattagt gcgatatttt ttggcgaata g 1551

```

<210> 63

<211> 914

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 63

```

Met Thr Asp Gln Arg Gly Pro Pro Pro Pro His Pro Gln Gln Ala Asn
1          5          10          15

```

```

Gly Tyr Lys Lys Phe Pro Pro His Asp Asn Gln Tyr Ser Gly Ala Asn
          20          25          30

```

```

Asn Ser Gln Pro Asn Asn His Tyr Asn Glu Asn Leu Tyr Ser Ala Arg
          35          40          45

```

```

Glu Pro His Asn Asn Lys Gln Tyr Gln Ser Lys Asn Gly Lys Tyr Gly
          50          55          60

```

```

Thr Asn Lys Tyr Asn Asn Arg Asn Asn Ser Gln Gly Asn Ala Gln Tyr
65          70          75          80

```

```

Tyr Asn Asn Arg Phe Asn Asn Gly Tyr Arg Leu Asn Asn Asn Asp Tyr
          85          90          95

```

```

Asn Pro Ala Met Leu Pro Gly Met Gln Trp Pro Ala Asn Tyr Tyr Ala
          100          105          110

```

```

Pro Gln Met Tyr Tyr Ile Pro Gln Gln Met Val Pro Val Ala Ser Pro
          115          120          125

```

```

Pro Tyr Thr His Gln Pro Leu Asn Thr Asn Pro Glu Pro Pro Ser Thr
          130          135          140

```

```

Pro Lys Thr Thr Lys Ile Glu Ile Thr Thr Lys Thr Gly Glu Arg Leu
          145          150          155          160

```

Asn Leu Lys Lys Phe His Glu Glu Lys Lys Ala Ser Lys Gly Glu Glu
 165 170 175

Lys Asn Asp Gly Val Glu Gln Lys Ser Lys Ser Gly Thr Pro Phe Glu
 180 185 190

Lys Glu Ala Thr Pro Val Leu Pro Ala Asn Glu Ala Val Lys Asp Thr
 195 200 205

Leu Thr Glu Thr Ser Asn Glu Lys Ser Thr Ser Glu Ala Glu Asn Thr
 210 215 220

Lys Arg Leu Phe Leu Glu Gln Val Arg Leu Arg Lys Ala Ala Met Glu
 225 230 235 240

Arg Lys Lys Asn Gly Leu Ile Ser Glu Thr Glu Lys Lys Gln Glu Thr
 245 250 255

Ser Asn His Asp Asn Thr Asp Thr Thr Lys Pro Asn Ser Val Ile Glu
 260 265 270

Ser Glu Pro Ile Lys Glu Ala Pro Lys Pro Thr Gly Glu Ala Asn Glu
 275 280 285

Val Val Ile Asp Gly Lys Ser Gly Ala Ser Val Lys Thr Pro Gln His
 290 295 300

Val Thr Gly Ser Val Thr Lys Ser Val Thr Phe Asn Glu Pro Glu Asn
 305 310 315 320

Glu Ser Ser Ser Gln Asp Val Asp Glu Leu Val Lys Asp Asp Asp Thr
 325 330 335

Thr Glu Ile Ser Asp Thr Thr Gly Gly Lys Thr Val Asn Lys Ser Asp
 340 345 350

Asp Glu Thr Ile Asn Ser Val Ile Thr Thr Glu Glu Asn Thr Val Lys
 355 360 365

Glu Thr Glu Pro Ser Thr Ser Asp Ile Glu Met Pro Thr Val Ser Gln
 370 375 380

Leu Leu Glu Thr Leu Gly Lys Ala Gln Pro Ile Ser Asp Ile Tyr Glu
 385 390 395 400

Phe Ala Tyr Pro Glu Asn Val Glu Arg Pro Asp Ile Lys Tyr Lys Lys
 405 410 415

Pro Ser Val Lys Tyr Thr Tyr Gly Pro Thr Phe Leu Leu Gln Phe Lys
 420 425 430

Asp Lys Leu Lys Phe Arg Pro Asp Pro Ala Trp Val Glu Ala Val Ser
 435 440 445

Ser Lys Ile Val Ile Pro Pro His Ile Ala Arg Asn Lys Pro Lys Asp
 450 455 460

Ser Gly Arg Phe Gly Gly Asp Phe Arg Ser Pro Ser Met Arg Gly Met
 465 470 475 480

Asp His Thr Ser Ser Ser Arg Val Ser Ser Lys Arg Arg Ser Lys Arg
 485 490 495

Met Gly Asp Asp Arg Arg Ser Asn Arg Gly Tyr Thr Ser Arg Lys Asp
 500 505 510

Arg Glu Lys Ala Ala Glu Lys Ala Glu Glu Gln Ala Pro Lys Glu Glu
 515 520 525

Ile Ala Pro Leu Val Pro Ser Ala Asn Arg Trp Ile Pro Lys Ser Arg
 530 535 540

Val Lys Lys Thr Glu Lys Lys Leu Ala Pro Asp Gly Lys Thr Glu Leu
 545 550 555 560

Phe Asp Lys Glu Glu Val Glu Arg Lys Met Lys Ser Leu Leu Asn Lys
 565 570 575

Leu Thr Leu Glu Met Phe Asp Ser Ile Ser Ser Glu Ile Leu Asp Ile
 580 585 590

Ala Asn Gln Ser Lys Trp Glu Asp Asp Gly Glu Thr Leu Lys Ile Val
 595 600 605

Ile Glu Gln Ile Phe His Lys Ala Cys Asp Glu Pro His Trp Ser Ser
 610 615 620

Met Tyr Ala Gln Leu Cys Gly Lys Val Val Lys Asp Leu Asp Pro Asn
 625 630 635 640

Ile Lys Asp Lys Glu Asn Glu Gly Lys Asn Gly Pro Lys Leu Val Leu
 645 650 655

His Tyr Leu Val Ala Arg Cys His Glu Glu Phe Glu Lys Gly Trp Ala

660	665	670
Asp Lys Leu Pro Ala Gly Glu Asp Gly Asn Pro Leu Glu Pro Glu Met 675 680 685		
Met Ser Asp Glu Tyr Tyr Ile Ala Ala Ala Ala Lys Arg Arg Gly Leu 690 695 700		
Gly Leu Val Arg Phe Ile Gly Tyr Leu Tyr Cys Leu Asn Leu Leu Thr 705 710 715 720		
Gly Lys Met Met Phe Glu Cys Phe Arg Arg Leu Met Lys Asp Leu Asn 725 730 735		
Asn Asp Pro Ser Glu Glu Thr Leu Glu Ser Val Ile Glu Leu Leu Asn 740 745 750		
Thr Val Gly Glu Gln Phe Glu His Asp Lys Phe Val Thr Pro Gln Ala 755 760 765		
Thr Leu Glu Gly Ser Val Leu Leu Asp Asn Leu Phe Met Leu Leu Gln 770 775 780		
His Ile Ile Asp Gly Gly Thr Ile Ser Asn Arg Ile Lys Phe Lys Leu 785 790 795 800		
Ile Asp Val Lys Glu Leu Arg Glu Ile Lys His Trp Asn Ser Ala Lys 805 810 815		
Lys Asp Ala Gly Pro Lys Thr Ile Gln Gln Ile His Gln Glu Glu Glu 820 825 830		
Gln Leu Arg Gln Lys Lys Asn Ser Gln Arg Ser Asn Ser Arg Phe Asn 835 840 845		
Asn His Asn Gln Ser Asn Ser Asn Arg Tyr Ser Ser Asn Arg Arg Asn 850 855 860		
Met Gln Asn Thr Gln Arg Asp Ser Phe Ala Ser Thr Lys Thr Gly Ser 865 870 875 880		
Phe Arg Asn Asn Gln Arg Asn Ala Arg Lys Val Glu Glu Val Ser Gln 885 890 895		
Ala Pro Arg Ala Asn Met Phe Asp Ala Leu Met Asn Asn Asp Gly Asp 900 905 910		

Ser Asp

<210> 64

<211> 2000

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 64

atgactgacc aaagaggacc accgccccca caccgcgagc aagccaatgg ctacaagaaa	60
tttcctcctc atgataacca atactctgga gccataata gtcagccaaa taaccactac	120
aatgaaaatc ttacagtgcc aagggaacct cacaataaca agcaatacca gtcgaaaaat	180
ggcaaatcgc ggacaaataa atacaataat cgtaataata gtcaaggaaa tgcacagtac	240
tacaataaca gattcaataa tggctacaga ctgaataaca acgactataa cccggcaatg	300
ttgccaggta tgcaatggcc agctaactac tatgctcctc agatgtatta tataccccaa	360
caaatgggtc cagttgcttc tccaccatat acacaccaac cactcaacac taatcccgaa	420
cctccctcta cacctaaaac tacaaaaata gaaatcacta cgaaaactgg tgaacgttta	480
aatttgaaaa aattccacga agaaaagaaa gcttcgaaaag gtgaagagaa gaatgatggg	540
gtagagcaga aatctaaatc aggaaccctt tttgaaaagg aggcaactcc tgtactacct	600
gctaataaag cagttaaaga tacgttaacc gaaacatcta atgaaaaatc tacctctgaa	660
gcggaaaata ctaagagatt attttttagag cagggtgagat tgcgcaaagc tgccatggaa	720
agaaagaaaa acggtttaat ctccgaaact gagaagaagc aagaaacttc aaatcatgat	780
aatactgata caactaagcc caactcgggt attgaatctg aacctattaa ggaagcgccg	840
aaaccaacgg gagaagctaa tgaagttgtg attgatggaa aatcaggggc aagtgtcaaa	900
actccacaac atgtaactgg aagtgttaact aaatccgtga cttttaacga acccgagaat	960
gagagcagtt ctccagatgt tgatgagctt gttaaagatg atgatactac tgaaatttcg	1020
gatacgactg gtggaaaaac tgtaataaaa agtgacgacg aaacaataaa ttccgtaatc	1080
accacagagg agaacacagt gaaggagaca gaaccttcca cttccgatat tgagatgcca	1140
actgtatctc aactacttga aactttaggt aaggctcaac caatttctga catatatgag	1200
tttgccctatc cggaaaatgt tgaaaggcct gatatacaat acaagaaacc aagcgttaag	1260
tacacttatg gccctacctt cttactgcaa ttcaaagata aactaaaatt caggcctgat	1320
cctgcgtggg ttgaagctgt atcctcgaaa attgttatac ctccctatat agccagaaat	1380
aaacccaaaag atagtggcag atttggaggc gatttcagaa gtccatctat gcgcggtatg	1440
gaccatactt ccagttcaag agtttcatcg aaaagaagat caaagagaat ggggtgacgac	1500
agaagatcta atagaggata tacctccaga aaagatcgcg aaaaagctgc agaaaaggca	1560
gaagagcaag ctccaaagga agaaatcgcc ccgttggttc cgagtgctaa tagatggata	1620

cctaaatcaa gggttaaaaa aacagaaaag aagtttagctc ctgacggcaa aactgaatta 1680
 ttgacaagg aagaagtaga acgaaaaatg aagtctctat taaacaagtt aacttttagaa 1740
 atgtttgatt caatctcctc tgagattttg gatatagcaa accaatcaaa gtgggaagat 1800
 gatggcgaaa cattgaaaat agttattgaa caaattttcc acaaggcttg tgatgaacct 1860
 cattggtctt caatgtatgc tcagttatgt ggcaaagtgg tcaaagacct tgacccaaac 1920
 attaaagata aagagaacga aggaaagaat ggaccaaagc ttgtgttgca ttatttagtg 1980
 gcaagatgtc acgaagaatt 2000

<210> 65

<211> 680

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 65

Met Thr Gln Phe Thr Asp Ile Asp Lys Leu Ala Val Ser Thr Ile Arg
 1 5 10 15

Ile Leu Ala Val Asp Thr Val Ser Lys Ala Asn Ser Gly His Pro Gly
 20 25 30

Ala Pro Leu Gly Met Ala Pro Ala Ala His Val Leu Trp Ser Gln Met
 35 40 45

Arg Met Asn Pro Thr Asn Pro Asp Trp Ile Asn Arg Asp Arg Phe Val
 50 55 60

Leu Ser Asn Gly His Ala Val Ala Leu Leu Tyr Ser Met Leu His Leu
 65 70 75 80

Thr Gly Tyr Asp Leu Ser Ile Glu Asp Leu Lys Gln Phe Arg Gln Leu
 85 90 95

Gly Ser Arg Thr Pro Gly His Pro Glu Phe Glu Leu Pro Gly Val Glu
 100 105 110

Val Thr Thr Gly Pro Leu Gly Gln Gly Ile Ser Asn Ala Val Gly Met
 115 120 125

Ala Met Ala Gln Ala Asn Leu Ala Ala Thr Tyr Asn Lys Pro Gly Phe
 130 135 140

Thr Leu Ser Asp Asn Tyr Thr Tyr Val Phe Leu Gly Asp Gly Cys Leu
 145 150 155 160

Gln Glu Gly Ile Ser Ser Glu Ala Ser Ser Leu Ala Gly His Leu Lys
 165 170 175

Leu Gly Asn Leu Ile Ala Ile Tyr Asp Asp Asn Lys Ile Thr Ile Asp
 180 185 190

Gly Ala Thr Ser Ile Ser Phe Asp Glu Asp Val Ala Lys Arg Tyr Glu
 195 200 205

Ala Tyr Gly Trp Glu Val Leu Tyr Val Glu Asn Gly Asn Glu Asp Leu
 210 215 220

Ala Gly Ile Ala Lys Ala Ile Ala Gln Ala Lys Leu Ser Lys Asp Lys
 225 230 235 240

Pro Thr Leu Ile Lys Met Thr Thr Thr Ile Gly Tyr Gly Ser Leu His
 245 250 255

Ala Gly Ser His Ser Val His Gly Ala Pro Leu Lys Ala Asp Asp Val
 260 265 270

Lys Gln Leu Lys Ser Lys Phe Gly Phe Asn Pro Asp Lys Ser Phe Val
 275 280 285

Val Pro Gln Glu Val Tyr Asp His Tyr Gln Lys Thr Ile Leu Lys Pro
 290 295 300

Gly Val Glu Ala Asn Asn Lys Trp Asn Lys Leu Phe Ser Glu Tyr Gln
 305 310 315 320

Lys Lys Phe Pro Glu Leu Gly Ala Glu Leu Ala Arg Arg Leu Ser Gly
 325 330 335

Gln Leu Pro Ala Asn Trp Glu Ser Lys Leu Pro Thr Tyr Thr Ala Lys
 340 345 350

Asp Ser Ala Val Ala Thr Arg Lys Leu Ser Glu Thr Val Leu Glu Asp
 355 360 365

Val Tyr Asn Gln Leu Pro Glu Leu Ile Gly Gly Ser Ala Asp Leu Thr
 370 375 380

Pro Ser Asn Leu Thr Arg Trp Lys Glu Ala Leu Asp Phe Gln Pro Pro
 385 390 395 400

Ser Ser Gly Ser Gly Asn Tyr Ser Gly Arg Tyr Ile Arg Tyr Gly Ile
 405 410 415

Arg Glu His Ala Met Gly Ala Ile Met Asn Gly Ile Ser Ala Phe Gly
 420 425 430

Ala Asn Tyr Lys Pro Tyr Gly Gly Thr Phe Leu Asn Phe Val Ser Tyr
 435 440 445

Ala Ala Gly Ala Val Arg Leu Ser Ala Leu Ser Gly His Pro Val Ile
 450 455 460

Trp Val Ala Thr His Asp Ser Ile Gly Val Gly Glu Asp Gly Pro Thr
 465 470 475 480

His Gln Pro Ile Glu Thr Leu Ala His Phe Arg Ser Leu Pro Asn Ile
 485 490 495

Gln Val Trp Arg Pro Ala Asp Gly Asn Glu Val Ser Ala Ala Tyr Lys
 500 505 510

Asn Ser Leu Glu Ser Lys His Thr Pro Ser Ile Ile Ala Leu Ser Arg
 515 520 525

Gln Asn Leu Pro Gln Leu Glu Gly Ser Ser Ile Glu Ser Ala Ser Lys
 530 535 540

Gly Gly Tyr Val Leu Gln Asp Val Ala Asn Pro Asp Ile Ile Leu Val
 545 550 555 560

Ala Thr Gly Ser Glu Val Ser Leu Ser Val Glu Ala Ala Lys Thr Leu
 565 570 575

Ala Ala Lys Asn Ile Lys Ala Arg Val Val Ser Leu Pro Asp Phe Phe
 580 585 590

Thr Phe Asp Lys Gln Pro Leu Glu Tyr Arg Leu Ser Val Leu Pro Asp
 595 600 605

Asn Val Pro Ile Met Ser Val Glu Val Leu Ala Thr Thr Cys Trp Gly
 610 615 620

Lys Tyr Ala His Gln Ser Phe Gly Ile Asp Arg Phe Gly Ala Ser Gly
 625 630 635 640

Lys Ala Pro Glu Val Phe Lys Phe Phe Gly Phe Thr Pro Glu Gly Val
 645 650 655

Ala Glu Arg Ala Gln Lys Thr Ile Ala Phe Tyr Lys Gly Asp Lys Leu
 660 665 670

Ile Ser Pro Leu Lys Lys Ala Phe
675 680

<210> 66

<211> 2000

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 66

atgactcaat tcactgacat tgataagcta gccgtctcca ccataagaat tttggctgtg	60
gacaccgtat ccaaggccaa ctcagggtcac ccagggtgtc cattgggtat ggcaccagct	120
gcacacgttc tatggagtca aatgcgcatg aacccaacca acccagactg gatcaacaga	180
gatagatttg tcttgtctaa cgggtcacgcg gtcgctttgt tgtattctat gctacatttg	240
actggttacg atctgtctat tgaagacttg aaacagttca gacagttggg ttccagaaca	300
ccagggtcac ctgaatttga gttgccaggt gttgaagtta ctaccggtcc attaggtcaa	360
ggtatctcca acgctgttgg tatggccatg gctcaagcta acctggctgc cacttacaac	420
aagccggggt ttacctgtgc tgacaactac acctatgttt tcttgggtga cggttgtttg	480
caagaaggta tttcttcaga agcttcctcc ttggctggtc atttgaaatt gggtaacttg	540
attgccatct acgatgacaa caagatcact atcgatggtg ctaccagtat ctcattcgat	600
gaagatgttg ctaagagata cgaagcctac ggttgggaag ttttgtacgt agaaaatggg	660
aacgaagatc tagccggtat tgccaaggct attgctcaag ctaagttatc caaggacaaa	720
ccaactttga tcaaatgac cacaaccatt gggtacggtt ccttgcacgc cggctctcac	780
tctgtgcacg gtgccccatt gaaagcagat gatgttaaac aactaaagag caaattcggg	840
ttcaaccag acaagtcctt tgttgttcca caagaagttt acgaccacta ccaaagaca	900
attttaaagc cagggtgtcg agccaacaac aagtggaaca agttgttcag cgaataccaa	960
aagaaattcc cagaattagg tgctgaattg gctagaagat tgagcggcca actaccgca	1020
aattgggaat ctaagttgcc aacttacacc gccaaaggact ctgccgtggc cactagaaaa	1080
ttatcagaaa ctgttcttga ggatgtttac aatcaattgc cagagttgat tgggtgttct	1140
gccgatttaa caccttctaa cttgaccaga tggaagggaag cccttgactt ccaacctcct	1200
tcttcgggtt caggtaacta ctctggtaga tacattagggt acggtattag agaacacgct	1260
atgggtgcca taatgaacgg tatttcagct ttcggtgcca actacaaacc atacggtggg	1320
actttcttga acttcgtttc ttatgtgtgt ggtgccgtta gattgtccgc tttgtctggc	1380
caccagttta tttgggttgc tacacatgac tctatcgggtg tcggtgaaga tgggtccaaca	1440
catcaacctt ttgaaacttt agcacacttc agatccctac caaacattca agtttggaga	1500
ccagctgatg gtaacgaagt ttctgccgcc tacaagaact ctttagaatc caagcatact	1560

```

ccaagtatca ttgctttgtc cagacaaaac ttgccacaat tggaaggtag ctctattgaa 1620
agcgcttcta aggggtggta cgtactacaa gatgttgcta acccagatat tatttttagtg 1680
gctactgggt ccgaagtgtc tttgagtgtt gaagctgcta agactttggc cgcaaagaac 1740
atcaaggctc gtgttggttc tctaccagat ttcttcactt ttgacaaaca acccctagaa 1800
tacagactat cagtcttacc agacaacgtt ccaatcatgt ctgttgaagt tttggctacc 1860
acatgttggg gcaaatacgc tcatcaatcc ttcgggtattg acagatttgg tgccctcggg 1920
aaggcaccag aagtcttcaa gttcttcggg ttcaccccag aagggtgttc tgaaagagct 1980
caaaagacca ttgcattcta 2000

```

```

<210> 67
<211> 196
<212> PRT
<213> Saccharomyces cerevisiae

```

```

<400> 67

```

```

Met Val Ala Gln Val Gln Lys Gln Ala Pro Thr Phe Lys Lys Thr Ala
1          5          10          15

```

```

Val Val Asp Gly Val Phe Asp Glu Val Ser Leu Asp Lys Tyr Lys Gly
20          25          30

```

```

Lys Tyr Val Val Leu Ala Phe Ile Pro Leu Ala Phe Thr Phe Val Cys
35          40          45

```

```

Pro Thr Glu Ile Ile Ala Phe Ser Glu Ala Ala Lys Lys Phe Glu Glu
50          55          60

```

```

Gln Gly Ala Gln Val Leu Phe Ala Ser Thr Asp Ser Glu Tyr Ser Leu
65          70          75          80

```

```

Leu Ala Trp Thr Asn Ile Pro Arg Lys Glu Gly Gly Leu Gly Pro Ile
85          90          95

```

```

Asn Ile Pro Leu Leu Ala Asp Thr Asn His Ser Leu Ser Arg Asp Tyr
100         105         110

```

```

Gly Val Leu Ile Glu Glu Glu Gly Val Ala Leu Arg Gly Leu Phe Ile
115         120         125

```

```

Ile Asp Pro Lys Gly Val Ile Arg His Ile Thr Ile Asn Asp Leu Pro
130         135         140

```

```

Val Gly Arg Asn Val Asp Glu Ala Leu Arg Leu Val Glu Ala Phe Gln
145         150         155         160

```

Trp Thr Asp Lys Asn Gly Thr Val Leu Pro Cys Asn Trp Thr Pro Gly
 165 170 175

Ala Ala Thr Ile Lys Pro Thr Val Glu Asp Ser Lys Glu Tyr Phe Glu
 180 185 190

Ala Ala Asn Lys
 195

<210> 68
 <211> 591
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 68
 atggtcgctc aagttcaaaa gcaagctcca acttttaaga aaactgccgt cgtcgacggt 60
 gtctttgacg aagtctcctt ggacaaatac aagggttaagt acgttgtcct agcctttatt 120
 ccattggcct tcactttcgt ctgtccaacc gaaatcattg ctttctcaga agctgctaag 180
 aaattcgaag aacaaggcgc tcaagttctt ttcgcctcca ctgactccga atactccctt 240
 ttggcatgga ccaatatccc aagaaaggaa ggtggtttgg gcccaatcaa cattccattg 300
 ttggctgaca ccaaccactc tttgtccaga gactatggtg tcttgatcga agaagaaggt 360
 gtcgccttga gaggtttgtt catcatcgac ccaaagggtg tcattagaca catcaccatt 420
 aacgatttgc cagtcggtag aaacgttgac gaagccttga gattgggtga agccttccaa 480
 tggaccgaca agaacggtac tgtcttgcca tgtaactgga ctccaggtgc tgctaccatc 540
 aagccaaccg ttgaagactc caaggaatac ttogaagctg ccaacaaata a 591

<210> 69
 <211> 291
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 69

Met Asn Ser Ile Leu Asp Arg Asn Val Arg Ser Ser Glu Thr Thr Leu
 1 5 10 15

Ile Lys Pro Glu Ser Glu Phe Asp Asn Trp Leu Ser Asp Glu Asn Asp
 20 25 30

Gly Ala Ser His Ile Asn Val Asn Lys Asp Ser Ser Ser Val Leu Ser
 35 40 45

Ala Ser Ser Ser Thr Trp Phe Glu Pro Leu Glu Asn Ile Ile Ser Ser
 50 55 60

Ala Ser Ser Ser Ser Ile Gly Ser Pro Ile Glu Asp Gln Phe Ile Ser
65 70 75 80

Ser Asn Asn Glu Glu Ser Ala Leu Phe Pro Thr Asp Gln Phe Phe Ser
85 90 95

Asn Pro Ser Ser Tyr Ser His Ser Pro Glu Val Ser Ser Ser Ile Lys
100 105 110

Arg Glu Glu Asp Asp Asn Ala Leu Ser Leu Ala Asp Phe Glu Pro Ala
115 120 125

Ser Leu Gln Leu Met Pro Asn Met Ile Asn Thr Asp Asn Asn Asp Asp
130 135 140

Ser Thr Pro Leu Lys Asn Glu Ile Glu Leu Asn Asp Ser Phe Ile Lys
145 150 155 160

Thr Asn Leu Asp Ala Lys Glu Thr Lys Lys Arg Ala Pro Arg Lys Arg
165 170 175

Leu Thr Pro Phe Gln Lys Gln Ala His Asn Lys Ile Glu Lys Arg Tyr
180 185 190

Arg Ile Asn Ile Asn Thr Lys Ile Ala Arg Leu Gln Gln Ile Ile Pro
195 200 205

Trp Val Ala Ser Glu Gln Thr Ala Phe Glu Val Gly Asp Ser Val Lys
210 215 220

Lys Gln Asp Glu Asp Gly Ala Glu Thr Ala Ala Thr Thr Pro Leu Pro
225 230 235 240

Ser Ala Ala Ala Thr Ser Thr Lys Leu Asn Lys Ser Met Ile Leu Glu
245 250 255

Lys Ala Val Asp Tyr Ile Leu Tyr Leu Gln Asn Asn Glu Arg Leu Tyr
260 265 270

Glu Met Glu Val Gln Arg Leu Lys Ser Glu Ile Asp Thr Leu Lys Gln
275 280 285

Asp Gln Lys
290

<210> 70
<211> 876

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 70

```

atgaactcta ttttagacag aaatgtaga tctagcgaaa ctactttaat taaacctgaa      60
tctgaatttg ataattgggt gtcggatgaa aatgacggag ctagtcatat caacgtcaac      120
aaggactcct cgtcagttct ttctgcatct tcttccacat gggtcgaacc attggaaaac      180
attatctcct ctgcatccag ctccctcgata ggctctccaa tcgaagacca gtttatatct      240
tccaacaacg aggaatctgc tctttttcca acagatcagt ttttcagtaa tccttcctca      300
tactcgcatt ctcccagggt tagcagctcg ataaaaagag aagaggatga caatgccttt      360
tcgttggcag attttgaacc ggcttctttg caattaatgc ctaacatgat aaatactgat      420
aataatgacg atagtacccc acttaagaat gaaatcgagc taaacgactc gtttataaaa      480
acaaatctag atgctaagga aacgaaaaag agggctccaa gaaaaagact gacccccttc      540
caaaagcaag ctcaacaaca gattgaaaaa cgctacagaa taaacatcaa cacaaagatt      600
gcaagactgc agcagattat cccatgggta gcaagtgaac aaacagcttt cgaagtaggt      660
gattctgtaa aaaaacagga cgaagacggc gcagaaactg ccgtactac tcctcttcca      720
tctgccgctg ctacaagcac gaagctaaat aaaagcatga tcctagaaaa agctgttgac      780
tatattctat atctacaaaa taacgaacga ctatacgaaa tggaagttca aaggttgaaa      840
agtgaaatcg acactttgaa acaagaccaa aaataa                                876

```

<210> 71

<211> 362

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 71

```

Met Ile Asn Asn Pro Lys Val Asp Ser Val Ala Glu Lys Pro Lys Ala
1              5              10             15

```

```

Val Thr Ser Lys Gln Ser Glu Gln Ala Ala Ser Pro Glu Pro Thr Pro
                20              25             30

```

```

Ala Pro Pro Val Ser Arg Asn Gln Tyr Pro Ile Thr Phe Asn Leu Thr
          35              40             45

```

```

Ser Thr Ala Pro Phe His Leu His Asp Arg His Arg Tyr Leu Gln Glu
          50              55             60

```

```

Gln Asp Leu Tyr Lys Cys Ala Ser Arg Asp Ser Leu Ser Ser Leu Gln
65              70             75             80

```

```

Gln Leu Ala His Thr Pro Asn Gly Ser Thr Arg Lys Lys Tyr Ile Val

```

85	90	95
Glu Asp Gln Ser Pro Tyr Ser Ser Glu Asn Pro Val Ile Val Thr Ser		
100	105	110
Ser Tyr Asn His Thr Val Cys Thr Asn Tyr Leu Arg Pro Arg Met Gln		
115	120	125
Phe Thr Gly Tyr Gln Ile Ser Gly Tyr Lys Arg Tyr Gln Val Thr Val		
130	135	140
Asn Leu Lys Thr Val Asp Leu Pro Lys Lys Asp Cys Thr Ser Leu Ser		
145	150	155
Pro His Leu Ser Gly Phe Leu Ser Ile Arg Gly Leu Thr Asn Gln His		
165	170	175
Pro Glu Ile Ser Thr Tyr Phe Glu Ala Tyr Ala Val Asn His Lys Glu		
180	185	190
Leu Gly Phe Leu Ser Ser Ser Trp Lys Asp Glu Pro Val Leu Asn Glu		
195	200	205
Phe Lys Ala Thr Asp Gln Thr Asp Leu Glu His Trp Ile Asn Phe Pro		
210	215	220
Ser Phe Arg Gln Leu Phe Leu Met Ser Gln Lys Asn Gly Leu Asn Ser		
225	230	235
Thr Asp Asp Asn Gly Thr Thr Asn Ala Ala Lys Lys Leu Pro Pro Gln		
245	250	255
Gln Leu Pro Thr Thr Pro Ser Ala Asp Ala Gly Asn Ile Ser Arg Ile		
260	265	270
Phe Ser Gln Glu Lys Gln Phe Asp Asn Tyr Leu Asn Glu Arg Phe Ile		
275	280	285
Phe Met Lys Trp Lys Glu Lys Phe Leu Val Pro Asp Ala Leu Leu Met		
290	295	300
Glu Gly Val Asp Gly Ala Ser Tyr Asp Gly Phe Tyr Tyr Ile Val His		
305	310	315
Asp Gln Val Thr Gly Asn Ile Gln Gly Phe Tyr Tyr His Gln Asp Ala		
325	330	335

117/148

Glu Lys Phe Gln Gln Leu Glu Leu Val Pro Ser Leu Lys Asn Lys Val
 340 345 350

Glu Ser Ser Asp Cys Ser Phe Glu Phe Ala
 355 360

<210> 72

<211> 1089

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 72

atgatcaata atcctaaggt agacagtgtg gcggagaaac ccaaagctgt gacatcaaag 60
 cagtcggagc aagcggttcc gccagaacca acaccagccc cccctgtttc tagaaatcag 120
 tatccgatca cgttcaactt gacttcaacc gcaccctttc atcttcatga ccgtcatcgc 180
 tacttacaag agcaagatct ttacaagtgc gcttctaggg attcgttggtc ttccctgcag 240
 caactcgccc atacacctaa tgggtcgaca aggaagaaat atattgttga ggaccaatct 300
 ccctatagtt cggaaaatcc agtcattgtg acctcttcct ataaccatac ggtttgcaca 360
 aactacttaa gaccaagaat gcagtttaca ggggtaccaga tatcaggata caaacgctat 420
 caggtaacag ttaacttaaa gactgtagat ttgccaaaga aagattgcac gtcgctgtcc 480
 cctcatattat ctggattttt gtcgataaga ggactcacga accaacaccc ggaaatcagt 540
 acatattttg aagcctacgc ggtaaaccac aaggaattag ggttcttggtc ctctagctgg 600
 aaagatgaac ctgtttttaa cgaattcaaa gccacagacc aaacagactt agaacttg 660
 ataaatttcc cctcctttag acagcttttc ctgatgagcc aaaaaaacgg totcaactca 720
 actgacgaca atggcactac aaatgcagcc aagaagttgc ctccacagca gcttcctact 780
 acaccaagcg cagacgctgg taacatatca agaattttta gccaaagagaa acaatttgac 840
 aactacttga acgaacgggt tatatttatg aaatggaagg aaaaattttt ggtaccagat 900
 gccctattaa tggaagggtg agacggcgca tcttatgatg ggttttatta cattgtccat 960
 gatcaagtta ccgggaacat tcaagggttt tactatcatc aagatgctga aaagttccaa 1020
 cagctggaat tagtaccatc ttgaaaaat aaagtcgagt ccagtgattg ttcttttgag 1080
 tttgcttga 1089

<210> 73

<211> 822

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 73

Met Leu Glu Gly Thr Val Asp Tyr Asp Pro Leu Glu Asp Ile Thr Asn
 1 5 10 15

Ile Leu Phe Ser Lys Glu Ser Leu Asn Asn Ile Asp Glu Leu Ile Ser
 20 25 30

Ile Thr Arg Ser Tyr Lys Lys Gln Leu Gln Glu Asp Ile Leu Lys Glu
 35 40 45

Glu Asn Glu Leu Lys Glu His Pro Lys Asn Ser Ala Glu Ile Glu Ala
 50 55 60

Ser Leu Arg Lys Val Phe Gln Asp Phe Lys Glu Thr Gln Asp Val Ser
 65 70 75 80

Ala Ser Thr Glu Leu Thr Ile Ser Asn Leu Thr Glu Gly Ile Ser Tyr
 85 90 95

Leu Asp Ile Ala Lys Lys Asn Leu Thr His Ser Leu Thr Leu Phe Gln
 100 105 110

Asn Leu Lys Ile Leu Thr Asp Ser Tyr Ile Gln Cys Asn Glu Leu Leu
 115 120 125

Ser Gln Gly Ser Phe Lys Lys Met Val Ser Pro Tyr Lys Ile Met Cys
 130 135 140

Ser Leu Ala Glu Asn Thr Phe Ile Ser Tyr Lys Ser Leu Asp Glu Ile
 145 150 155 160

Asn Tyr Leu Leu Ser Ser Ile Ser Arg Leu Lys Gly Asp Thr Leu Ser
 165 170 175

Lys Ile Lys Gln Asn Tyr Asn Ala Leu Phe Ser Gly Gly Asn Ile Ser
 180 185 190

Glu His Asp Thr Ala Leu Thr Met Glu Leu Arg Glu Gly Ala Cys Glu
 195 200 205

Leu Leu Asp Cys Asp Thr Ser Thr Arg Ala Gln Met Ile Asp Trp Cys
 210 215 220

Leu Asp Lys Leu Leu Phe Glu Met Lys Glu Ile Phe Arg Val Asp Asp
 225 230 235 240

Glu Ala Gly Ser Leu Glu Asn Leu Ser Arg Arg Tyr Ile Tyr Phe Lys
 245 250 255

Lys Ile Leu Asn Asn Phe Asn Ser Lys Phe Ala Asp Tyr Phe Leu Lys
 260 265 270

Asp Trp Glu Met Ala Val Arg Leu Thr Thr Thr Phe Tyr His Ile Thr
 275 280 285

His Lys Asp Leu Gln Thr Leu Leu Lys Arg Glu Phe Lys Asp Lys Asn
 290 295 300

Pro Ser Ile Asp Leu Phe Met Thr Ala Leu Gln Ser Thr Leu Asp Phe
 305 310 315 320

Glu Lys Tyr Ile Asp Val Arg Phe Ser Lys Lys Ile Lys Glu Pro Lys
 325 330 335

Leu Ser Ser Cys Phe Glu Pro Tyr Leu Thr Leu Trp Val Ser His Gln
 340 345 350

Asn Gln Met Met Glu Lys Lys Phe Leu Ser Tyr Met Ser Glu Pro Lys
 355 360 365

Tyr Pro Ser Asn Glu Thr Glu Ser Leu Val Leu Pro Ser Ser Ala Asp
 370 375 380

Leu Phe Arg Thr Tyr Arg Ser Val Leu Thr Gln Thr Leu Glu Leu Ile
 385 390 395 400

Asp Asn Asn Ala Asn Asp Ser Ile Leu Thr Ser Leu Ala Asn Phe Phe
 405 410 415

Ser Arg Trp Leu Gln Thr Tyr Ser Gln Lys Ile Leu Leu Pro Leu Leu
 420 425 430

Leu Pro Asp Asn Ile Glu Val Gln Asp Lys Leu Glu Ala Ala Lys Tyr
 435 440 445

Thr Val Leu Leu Ile Asn Thr Ala Asp Tyr Cys Ala Thr Thr Ile Asp
 450 455 460

Gln Leu Glu Asp Lys Leu Ser Glu Phe Ser Gly Asn Arg Glu Lys Leu
 465 470 475 480

Ala Asn Ser Phe Thr Lys Thr Lys Asn Ile Tyr Asp Asp Leu Leu Ala
 485 490 495

Lys Gly Thr Ser Phe Leu Leu Asn Arg Val Ile Pro Leu Asp Leu Asn
 500 505 510

Phe Val Trp Arg Glu Phe Ile Asn Asn Asp Trp Ser Asn Ala Ala Ile

515		520		525
Glu Asp Tyr Ser Arg Tyr Met Val Thr Leu Lys Ser Val Leu Lys Met				
530		535		540
Pro Ala Leu Thr Asp Ala Ser Ile Lys Gln Gln Gln Glu Gln Pro Ser				
545		550		555
Thr Leu Ala Phe Ile Leu Ser Gln Phe Asn Arg Asp Val Tyr Lys Trp				
	565		570	575
Asn Phe Leu Asp Lys Val Ile Asp Ile Ile Thr Thr Asn Phe Val Ser				
	580		585	590
Asn Thr Ile Arg Leu Leu Gln Pro Val Pro Pro Phe Ser Leu Ala Gly				
	595		600	605
Ser Lys Arg Lys Phe Glu Thr Arg Thr Val Val Asn Ile Gly Glu Gln				
610		615		620
Leu Leu Leu Asp Leu Glu Leu Leu Lys Glu Ile Phe His Thr Leu Pro				
625		630		635
Glu Ser Val Ser Asn Asp Ser Asp Leu Arg Glu Asn Thr Ser Tyr Lys				
	645		650	655
Arg Val Lys Arg His Ala Asp Asn Asn Ile Asp Gln Leu Leu Lys Phe				
	660		665	670
Ile Lys Leu Leu Met Ala Pro Leu Asp Ser Ala Asp Asp Tyr Tyr Glu				
	675		680	685
Thr Tyr Ser Lys Leu Thr Asn Asn Asn Pro Asp Ser Ala Val Trp Ser				
690		695		700
Phe Val Leu Ala Leu Lys Gly Ile Pro Trp Asp Leu Ala Leu Trp Lys				
705		710		715
Lys Leu Trp Ser Ala Tyr Asn Leu Glu Thr Asp Asp Thr Asp Glu Gly				
	725		730	735
Ser Arg Pro Asp Ser Asn Arg Asp Leu Phe Ile Phe Lys Trp Asp Lys				
	740		745	750
Val Leu Leu Gly Gln Phe Glu Asn Asn Leu Ala Arg Met Gln Asp Pro				
	755		760	765

121/148

Asn Trp Ser Lys Phe Val Arg Gln Asp Leu Lys Ile Ser Pro Pro Val
 770 775 780

Met Lys Arg Ile Val Ser Thr Pro Gln Ile Gln Gln Gln Lys Glu Glu
 785 790 795 800

Gln Lys Lys Gln Ser Leu Ser Val Lys Asp Phe Val Ser His Ser Arg
 805 810 815

Phe Phe Asn Arg Gly Thr
 820

<210> 74

<211> 2000

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 74

atgctggaag gtacggtaga ttatgacccg ctggaagata ttaccaatat actttttttca	60
aaagaatccc tgaacaacat agatgaactg atcagtatta ccagaagcta caaaaagcaa	120
ttgcaagagg atatttctcaa agaagagaat gaattgaagg aacaccctaa aaattccgct	180
gaaatagagg cttctctgag gaaagttttc caagatttca aagaaactca agatgtctca	240
gcctccaccg agttgacgat atcgaatctg acagaaggta tctcgtacct ggacattgcc	300
aagaaaaacc tcaccactc tttgactctt ttccaaaatt taaagatatt gacagacagt	360
tacatacaat gcaatgaatt actctcacag ggctcattca aaaaaatggg gtcccccttat	420
aagataatgt gttcgcttgc tgaaaacaca ttcattctctt acaaatcatt ggacgagata	480
aactatttgt tgagctccat ttcaagactg aaaggagaca ctttgtccaa aattaaacaa	540
aactacaatg cgctcttttc cggcggcaat atctcagagc atgatacagc actcactatg	600
gaattgcgcg aagggtgcctg cgagctactc gactgcgata caagtacgag agcccagatg	660
atagattggg gtttggacaa acttctcttc gaaatgaaag agatatttag ggtcgacgat	720
gaagccggat cctagaaaa tttatcgaga agatacattt acttcaaaaa aattcttaat	780
aacttcaatt caaagttcgc agactatttc ttaaaagact gggaaatggc agtcagattg	840
accaccactt tttatcacat tacacacaag gaccttcaga cacttctgaa aagggaattc	900
aaagacaaga acccttccat tgatctattc atgacagcat tacaatcgac gctagatttc	960
gaaaaataca tcgacgtacg attttcaaaa aaaattaagg aacccaaact aagttcctgc	1020
ttcgaacctt atttgacttt atgggtgtct caccaaaacc aaatgatgga aaagaaattt	1080
ctttcttata tgagtgagcc gaagtaccca tctaataaaa cagaatctct cgtgttacct	1140
tcgagtgcag accttttcag gacatatcgt tccgtactga ctcagacctt agagctcatt	1200
gataataatg ccaatgatag catattgact tcattggcaa attttttcag tagatggctt	1260

```

caaacttact cacaaaaaat tcttcttcct ttactgctgc ccgacaatat tgaagtccag 1320
gataagctag aagctgccaa gtataccggt ttattgatca atactgcaga ttattgtgcc 1380
acgactatag atcaattgga ggataaatta tctgaattca gcggtaatcg tgaaaagctg 1440
gcaaacagtt ttacgaaaac gaaaaatata tacgacgatt tactagcaaa aggaacttct 1500
tttctattaa accgtgtcat acccttagat ctaaattttg tatggagaga gtttatcaac 1560
aatgattggt caaatgctgc gatagaagat tatagcaggt acatggtaac cctcaaatcc 1620
gtacttaaaa tgcccgcatt aacagatgcc tctattaaac aacagcaaga gcaaccttcg 1680
actttggcat ttattttgtc gcaattcaat agagatgttt ataagtggaa tttcttgat 1740
aaggtgattg atatcatcac tacaaatttt gtaagcaata ccatccgcct tctgcagccc 1800
gttccacctt tttccctggc gggcagcaaa aggaaatttg aaaccagaac tgttgtcaac 1860
attggcgagc agcttctcct tgatttagaa ttgctgaagg agatttttca cactttacca 1920
gaaagtgtaa gtaacgattc tgacttgcca gaaaatacct cttataagag ggtgaaaaga 1980
catgcagaca ataatataga 2000

```

```

<210> 75
<211> 779
<212> PRT
<213> Saccharomyces cerevisiae

```

```

<400> 75

```

```

Met Glu Arg Thr Asn Thr Thr Thr Phe Lys Phe Phe Ser Leu Gly Gly
1           5           10          15

```

```

Ser Asn Glu Val Gly Arg Ser Cys His Ile Leu Gln Tyr Lys Gly Lys
20          25          30

```

```

Thr Val Met Leu Asp Ala Gly Ile His Pro Ala Tyr Gln Gly Leu Ala
35          40          45

```

```

Ser Leu Pro Phe Tyr Asp Glu Phe Asp Leu Ser Lys Val Asp Ile Leu
50          55          60

```

```

Leu Ile Ser His Phe His Leu Asp His Ala Ala Ser Leu Pro Tyr Val
65          70          75          80

```

```

Met Gln Arg Thr Asn Phe Gln Gly Arg Val Phe Met Thr His Pro Thr
85          90          95

```

```

Lys Ala Ile Tyr Arg Trp Leu Leu Arg Asp Phe Val Arg Val Thr Ser
100         105         110

```

123/148

Ile Gly Ser Ser Ser Ser Ser Met Gly Thr Lys Asp Glu Gly Leu Phe
 115 120 125

Ser Asp Glu Asp Leu Val Asp Ser Phe Asp Lys Ile Glu Thr Val Asp
 130 135 140

Tyr His Ser Thr Val Asp Val Asn Gly Ile Lys Phe Thr Ala Phe His
 145 150 155 160

Ala Gly His Val Leu Gly Ala Ala Met Phe Gln Ile Glu Ile Ala Gly
 165 170 175

Leu Arg Val Leu Phe Thr Gly Asp Tyr Ser Arg Glu Val Asp Arg His
 180 185 190

Leu Asn Ser Ala Glu Val Pro Pro Leu Ser Ser Asn Val Leu Ile Val
 195 200 205

Glu Ser Thr Phe Gly Thr Ala Thr His Glu Pro Arg Leu Asn Arg Glu
 210 215 220

Arg Lys Leu Thr Gln Leu Ile His Ser Thr Val Met Arg Gly Gly Arg
 225 230 235 240

Val Leu Leu Pro Val Phe Ala Leu Gly Arg Ala Gln Glu Ile Met Leu
 245 250 255

Ile Leu Asp Glu Tyr Trp Ser Gln His Ala Asp Glu Leu Gly Gly Gly
 260 265 270

Gln Val Pro Ile Phe Tyr Ala Ser Asn Leu Ala Lys Lys Cys Met Ser
 275 280 285

Val Phe Gln Thr Tyr Val Asn Met Met Asn Asp Asp Ile Arg Lys Lys
 290 295 300

Phe Arg Asp Ser Gln Thr Asn Pro Phe Ile Phe Lys Asn Ile Ser Tyr
 305 310 315 320

Leu Arg Asn Leu Glu Asp Phe Gln Asp Phe Gly Pro Ser Val Met Leu
 325 330 335

Ala Ser Pro Gly Met Leu Gln Ser Gly Leu Ser Arg Asp Leu Leu Glu
 340 345 350

Arg Trp Cys Pro Glu Asp Lys Asn Leu Val Leu Ile Thr Gly Tyr Ser
 355 360 365

Ile Glu Gly Thr Met Ala Lys Phe Ile Met Leu Glu Pro Asp Thr Ile
 370 375 380

Pro Ser Ile Asn Asn Pro Glu Ile Thr Ile Pro Arg Arg Cys Gln Val
 385 390 395 400

Glu Glu Ile Ser Phe Ala Ala His Val Asp Phe Gln Glu Asn Leu Glu
 405 410 415

Phe Ile Glu Lys Ile Ser Ala Pro Asn Ile Ile Leu Val His Gly Glu
 420 425 430

Ala Asn Pro Met Gly Arg Leu Lys Ser Ala Leu Leu Ser Asn Phe Ala
 435 440 445

Ser Leu Lys Gly Thr Asp Asn Glu Val His Val Phe Asn Pro Arg Asn
 450 455 460

Cys Val Glu Val Asp Leu Glu Phe Gln Gly Val Lys Val Ala Lys Ala
 465 470 475 480

Val Gly Asn Ile Val Asn Glu Ile Tyr Lys Glu Glu Asn Val Glu Ile
 485 490 495

Lys Glu Glu Ile Ala Ala Lys Ile Glu Pro Ile Lys Glu Glu Asn Glu
 500 505 510

Asp Asn Leu Asp Ser Gln Ala Glu Lys Gly Leu Val Asp Glu Glu Glu
 515 520 525

His Lys Asp Ile Val Val Ser Gly Ile Leu Val Ser Asp Asp Lys Asn
 530 535 540

Phe Glu Leu Asp Phe Leu Ser Leu Ser Asp Leu Arg Glu His His Pro
 545 550 555 560

Asp Leu Ser Thr Thr Ile Leu Arg Glu Arg Gln Ser Val Arg Val Asn
 565 570 575

Cys Lys Lys Glu Leu Ile Tyr Trp His Ile Leu Gln Met Phe Gly Glu
 580 585 590

Ala Glu Val Leu Gln Asp Asp Asp Arg Val Thr Asn Gln Glu Pro Lys
 595 600 605

Val Lys Glu Glu Ser Lys Asp Asn Leu Thr Asn Thr Gly Lys Leu Ile
 610 615 620

Leu Gln Ile Met Gly Asp Ile Lys Leu Thr Ile Val Asn Thr Leu Ala
625 630 635 640

Val Val Glu Trp Thr Gln Asp Leu Met Asn Asp Thr Val Ala Asp Ser
645 650 655

Ile Ile Ala Ile Leu Met Asn Val Asp Ser Ala Pro Ala Ser Val Lys
660 665 670

Leu Ser Ser His Ser Cys Asp Asp His Asp His Asn Asn Val Gln Ser
675 680 685

Asn Ala Gln Gly Lys Ile Asp Glu Val Glu Arg Val Lys Gln Ile Ser
690 695 700

Arg Leu Phe Lys Glu Gln Phe Gly Asp Cys Phe Thr Leu Phe Leu Asn
705 710 715 720

Lys Asp Glu Tyr Ala Ser Asn Lys Glu Glu Thr Ile Thr Gly Val Val
725 730 735

Thr Ile Gly Lys Ser Thr Ala Lys Ile Asp Phe Asn Asn Met Lys Ile
740 745 750

Leu Glu Cys Asn Ser Asn Pro Leu Lys Gly Arg Val Glu Ser Leu Leu
755 760 765

Asn Ile Gly Gly Asn Leu Val Thr Pro Leu Cys
770 775

<210> 76

<211> 2000

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 76

atggagcgaa caaatacaac aacattcaaa tttttttcat tgggaggaag taacgaagtt 60

ggacgatcat gtcatatatt gcaatataaa ggtaaaacag tgatgctcga cgcaggaatt 120

catccggcat atcaaggatt agcttcatta cctttttacg atgaatttga tctttccaaa 180

gtcgatatct tactgatatc acatttccat ttagaccatg ccgcttcact tccgtatgtg 240

atgcaacgga ctaactttca aggcagagtt ttcattgacac atccaaccaa agccatttac 300

agatggctac tgcgagattt tgtaagagtt actagtatag gttcttcattc ctctctatg 360

gggactaaag acgaagggtct attctcagat gaggatttag ttgattcttt cgataaaatt 420

gaaacggtgg actatcattc tactgttgac gtcaatggta tcaaatttac ggcatccat 480

```

gcaggccatg tattgggtgc agcgatgttt caaatagaga ttgctggcct tagggtgtta    540
ttcacagggtg actattcgag agaagtagat cgtcacttaa attctgctga ggtgcctcca    600
ctttcatcca acgtattaat tgtggaatct acctttggta cgcctactca cgagccccgt    660
ttaaatagag aaaggaagct aaccctaact atccattcca cagtgatgag aggaggccgt    720
gttctactgc ctgtttttgc tttagggaga gctcaagaaa tcatgcttat actggatgag    780
tactgggtctc aacatgctga tgaactcggg ggtggacaag tccaatatt ttatgcatca    840
aatttggtcaa aaaaatgtat gagcgttttt caaacctatg taaatatgat gaatgatgac    900
attagaaaaa aatttagaga ctctcagact aatcccttca tattcaaaaa tatatcttat    960
cttagaaaact tggaggattt ccaagatttt ggtcccagtg tgatgttggc ctcaccaggt   1020
atgctgcaaa gtgggttatc aagagattta cttgaaaggt ggtgccctga agataaaaaat   1080
ctagtactaa tcaccgggta ctccatcgaa ggaacaatgg cgaaatttat tatgcttgag   1140
ccagatacaa taccttccat aaataatccg gaaataacca ttccaagacg ttgtcaagtt   1200
gaagaaatct cctttgccgc acacgttgac ttccaggaaa atttagaatt tattgaaaag   1260
attagtgcgc caaatatcat ccttgttcat ggagaagcca atcccatggg ccgtttgaaa   1320
tctgcattgt tatccaattt cgcgtcttta aagggtacag ataatgaagt ccatgttttt   1380
aatcctagaa actgtgttga agtagatctt gaatttcaag gtgtcaaggt tgcaaaagct   1440
gtgggaaata ttgtgaacga aatatataaa gaagaaaacg tagagataaa ggaagaaatt   1500
gcggctaaaa ttgaacctat aaaggaagaa aacgaagaca acttggaattc tcaagcagaa   1560
aaaggtttgg ttgatgaaga ggaacacaaa gacatagtcg tttctgggat tctagtttca   1620
gatgacaaaa atttcgaatt agacttcctt tctttgtctg atttaagaga gcaccatccc   1680
gatctttcta caacgatatt aagagagcgc cagtcagttc gtgtaaattg taaaaaggag   1740
ctaatttact ggcacatttt acaaagtgtt ggagaggctg aagttcttca agatgatgat   1800
agagtaacaa atcaagaacc aaagggttaa gaagagtcaa aagacaatct gaccaataca   1860
ggtaaattga ttctacagat aatgggtgat attaagttaa ctattgttaa tactctagcc   1920
gttgtggaat ggactcaaga tttaatgaat gacactgtag cagactccat tattgctata   1980
cttatgaatg tggattcagc
                                                                 2000

```

<210> 77

<211> 208

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 77

Met Ser Leu Ile His Pro Asp Thr Ala Lys Tyr Pro Phe Lys Phe Glu
 1 5 10 15

Pro Phe Leu Arg Gln Glu Tyr Ser Phe Ser Leu Asp Pro Asp Arg Pro
 20 25 30

Ile Cys Glu Phe Tyr Asn Ser Arg Glu Gly Pro Lys Ser Cys Pro Arg
 35 40 45

Gly Pro Leu Cys Pro Lys Lys His Val Leu Pro Ile Phe Gln Asn Lys
 50 55 60

Ile Val Cys Arg His Trp Leu Arg Gly Leu Cys Lys Lys Asn Asp Gln
 65 70 75 80

Cys Glu Tyr Leu His Glu Tyr Asn Leu Arg Lys Met Pro Glu Cys Val
 85 90 95

Phe Phe Ser Lys Asn Gly Tyr Cys Thr Gln Ser Pro Asp Cys Gln Tyr
 100 105 110

Leu His Ile Asp Pro Ala Ser Lys Ile Pro Lys Cys Glu Asn Tyr Glu
 115 120 125

Met Gly Phe Cys Pro Leu Gly Ser Ser Cys Pro Arg Arg His Ile Lys
 130 135 140

Lys Val Phe Cys Gln Arg Tyr Met Thr Gly Phe Cys Pro Leu Gly Lys
 145 150 155 160

Asp Glu Cys Asp Met Glu His Pro Gln Phe Ile Ile Pro Asp Glu Gly
 165 170 175

Ser Lys Leu Arg Ile Lys Arg Asp Asp Glu Ile Asn Thr Arg Lys Met
 180 185 190

Asp Glu Glu Lys Glu Arg Arg Leu Asn Ala Ile Ile Asn Gly Glu Val
 195 200 205

<210> 78

<211> 627

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 78

atgagcctaa ttcaccccca tacagcaaaa taticcttta aatttgaacc tttcctcagg 60

caagagtatt cgttttcact cgatcctgac agacctatct gtgaatttta caattctaga 120

gaaggcccta aatcatgtcc gaggggaccg ttatgtccaa aaaagcatgt gttaccaata 180

tttcagaata aaattgtttg tagacattgg cttcgagggt tgtgcaaaaa gaatgaccaa 240

```

tgtgaatact tacatgaata caatcttcga aaaatgcctg aatgtgtctt cttcagcaaa   300
aacgggtact gtacacaaag tccagattgt caatatctac acatagatcc cgctagcaag   360
ataccaaaat gtgaaaatta cgaaatggga ttctgtcctc tggggagttc ttgtcctaga   420
cggcatatta agaaggtttt ctgtcaaaga tacatgaccg gattttgtcc tttagggaag   480
gatgaatgtg atatggaaca tccacagttc ataatcccag atgaaggtag taaattaaga   540
attaagagag acgatgagat aaataccagg aaaatggatg aagaaaagga aaggcgttta   600
aacgcaatta taaacgggtga agtttga                                     627

```

```

<210> 79
<211> 107
<212> PRT
<213> Saccharomyces cerevisiae

```

```

<400> 79

```

```

Met Val Lys Gly Lys Thr Phe Leu Lys Arg Ile Cys Pro Glu Glu Thr
1              5              10              15

```

```

Leu Asn Glu Glu Thr Lys Gln Glu Val Ser Val Gly Phe Asp Lys Met
                20              25              30

```

```

Arg Thr Leu Leu Arg Ser Arg Glu Ser Gly Met Thr Phe Ser Gln Gly
35              40              45

```

```

Pro Lys Leu Ala Ser Cys Gln Ser Val Ile Asn Ala Ser Ser Glu Lys
50              55              60

```

```

Thr Ala Trp Thr Gln Leu Val Phe Arg Lys Ser Lys Met Lys Thr Tyr
65              70              75              80

```

```

Thr Lys Ser Val His Val Ile Phe Ile Ala Met Gly Glu Gly Glu Asp
85              90              95

```

```

Glu Ser Val Asp Met Asn Val Gly Ile Ser Tyr
100              105

```

```

<210> 80
<211> 324
<212> DNA
<213> Saccharomyces cerevisiae

```

```

<400> 80

```

```

atggtaaagg gtaaaacggt tctgaaaaga atctgtccgg aagaaacggt aaacgaagaa   60

```

```

actaagcagg aagtttcggt agggttcgat aagatgagaa ccctgttgcg gtctcgagaa   120

```

```

tcagggatga ctttctccca aggacctaa ttagccagtt gccaatcagt gataaatgca   180

```

tcacatgaaa aaacggcttg gacacaactc gtgttttagga agagtaaaat gaagacgtac 240
 accaagtctg tacacgttat cttcattgct atgggggaag gggaggatga aagtgttgat 300
 atgaatgtag gtattagtta ttaa 324

<210> 81
 <211> 425
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 81

Met Thr Asp Pro Arg Arg Arg Thr Gly Arg His Phe Leu Thr Pro Glu
 1 5 10 15

Asn Leu Ser Ser Thr Leu Gln Ile Thr Asn Leu Pro Pro Glu Trp Asn
 20 25 30

Gln Asp Ile Ile Thr Ser Val Val Ala Gly Ser Gly Pro Val Ile Asp
 35 40 45

Ile Lys Ala Lys Asn Asp Pro Arg Thr Gly Lys Leu Thr Gly Val Leu
 50 55 60

Phe Asp Tyr Leu Thr Ser Lys Asp Cys Lys Arg Ala Trp Glu Ile Leu
 65 70 75 80

Asn Arg Ile Glu Asn Phe Pro Val Lys Ile Glu Gln Ile Ile Pro Pro
 85 90 95

Asn Tyr Lys Asp His Leu Arg Glu Thr Ala Asn Lys Asn Ser Gln Lys
 100 105 110

Gln Val Leu Gln Leu Asn Arg Asp Ser Tyr Pro Phe Glu Ala Gly Leu
 115 120 125

Glu Leu Pro Phe Glu Met Val Thr Glu Val Pro Ile Pro Arg Arg Pro
 130 135 140

Pro Pro Pro Gln Ala Ala Asn Asn Thr Asn Ser Val Ser Asn Asn Thr
 145 150 155 160

Asn Ile Gln Phe Pro Asp Ile Leu Ser Lys Ala Ser Lys His Leu Pro
 165 170 175

Ser Phe Gln Asp Gly Ser Ile Ile Ala Pro Asp Lys Ile Ser Gln Asn
 180 185 190

Leu Ser Lys Ile Pro Pro Leu Gln Leu Ile Glu Ile Ile Ser Asn Leu

195 200 205
 Lys Ile Leu Ser Asn Gln Glu Asn Ile Gln Lys Ser Gln Leu Glu Ser
 210 215 220
 Phe Leu Asp Thr Asn Ser Asp Ile Thr Ile Ser Val Thr Gln Ala Leu
 225 230 235 240
 Leu Glu Met Gly Phe Ile Asp Tyr Ser Val Val Thr Lys Val Leu Lys
 245 250 255
 Ser Gln Val Gly Glu Ala Pro Ser Leu Leu Ser Ser Asn Asn Thr Ser
 260 265 270
 Asn Ser Asn Thr Pro Val Ser Val Ile Arg Asn Asn Thr Pro Leu His
 275 280 285
 Val Pro Ser Asn Glu Val Ser Asn Asn Pro Asn Asn Met Pro Leu Asn
 290 295 300
 Val Ala Met Pro Met Pro Met Ser Thr Pro Pro Phe Ile Pro Leu Pro
 305 310 315 320
 Leu Gln Gln Gln Pro Phe Gly Phe Ala Pro Pro Gly Pro Phe Met Pro
 325 330 335
 Pro Ala Gln Gly Pro Ser Met Gly Gln Pro Val Leu Ala Asn Gln Leu
 340 345 350
 Gly Gln Val Gln Gln Gln Asn Ile Ser Ser Thr Glu Gly Pro Ser Asn
 355 360 365
 Ala Asn Lys Ala Asn Asp Ser Gly Thr Ile Asn Met Ala Lys Leu Gln
 370 375 380
 Leu Leu Pro Glu Asn Gln Gln Asp Met Ile Lys Gln Val Leu Thr Leu
 385 390 395 400
 Thr Pro Ala Gln Ile Gln Ser Leu Pro Ser Asp Gln Gln Leu Met Val
 405 410 415
 Glu Asn Phe Arg Lys Glu Tyr Ile Ile
 420 425

<210> 82

<211> 1278

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 82
 atgacagatc ccagaagaag aacaggccgt catttcttga caccggagaa tttatcctct 60
 acattacaaa tcacaaactt acctccagaa tggaaccaag atataattac ttcggtcgtg 120
 gccggttctg gtccagttat agatataaaa gctaagaatg acccgagAAC tggtaaacta 180
 accggtgtac tgttcgatta tttgactagt aaagattgta aacgcgcttg ggaaatttta 240
 aatagaattg aaaactttcc cgtaaagata gagcaaataa tcccaccaa ttataaggac 300
 catcttagag aaacagcaaa taaaaattct caaaagcagg tattacaact taatagagat 360
 tcgtaccctt tcgaggcggg tttggagcta cctttcgaaa tggtgacaga agtccccatt 420
 cctaggcgac caccgccacc acaggctgca aataacacaa actctgtatc aaataacaca 480
 aacattcaat tccccgacat actaagtaaa gcatctaaac acttgccaag tttccaagat 540
 ggctcgatta ttgcaccaga caaaatttca caaaatttaa gtaaaattcc gccgttgcaa 600
 cttattgaaa ttatatcaaa tttgaaaata ttatcaaacc aagaaaacat ccaaaaatcg 660
 caattagaat ctttcttaga tactaacagt gatatcacia tatcagtgc ccaagcccta 720
 ctgaaatgg gatttataga ctacagcgtg gtgactaaag tgttgaaatc ccaagttggc 780
 gagggcccat ctttgctttc gagtaataac acaagtaatt cgaacacccc cgtaagcgta 840
 attagaaata aactccgtt gcatgtacct tctaataag tcagcaacia tcctaacaat 900
 atgccactga acgtagctat gccaatgcct atgtcgacac caccatttat ccctttacct 960
 ctgcaacaac aaccgttcgg ttttgcgcca cggggccctt tcatgcctcc agtcaaggc 1020
 ccctccatgg gacagcctgt gttggcaaat caactcggcc aggtccagca acaaaatata 1080
 agttctacag aaggaccctc taacgcgaat aaagcgaatg acagcggcac cattaatatg 1140
 gcgaaactgc aattactacc tgaaaaccaa caagatatga tcaaacaagt tcttactttg 1200
 acacctgccc agatccaaag tttaccaagt gaccagcaac ttatggtgga aaactttaga 1260
 aaagaatata taatctaa 1278

<210> 83
 <211> 1592
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 83

Met Gly Thr Asp Pro Leu Ile Ile Arg Asn Asn Gly Ser Phe Trp Glu
 1 5 10 15

Val Asp Asp Phe Thr Arg Leu Gly Arg Thr Gln Leu Leu Ser Tyr Tyr
 20 25 30

Leu Pro Leu Ala Ile Ile Ala Ser Ile Gly Ile Phe Ala Leu Cys Arg

35	40	45	
Ser Gly Leu Ser Arg Tyr Val Arg Ser Ala Glu Cys Asp Leu Val Asn			
50	55	60	
Glu Tyr Leu Phe Gly Ala Gln Glu Glu Arg Lys Glu Asp Asn Ser Ile			
65	70	75	80
Glu Arg Leu Leu Arg Asn Ser Asn Thr Gln Ala Asn Tyr Val Asn Val			
85	90	95	
Lys Lys Gln Gly Arg Ile Leu Lys Leu Arg His Phe Asp Ile Thr Thr			
100	105	110	
Ile Asp Val Lys Gln Ile Asp Ala Lys Asn His Gly Gly Leu Thr Phe			
115	120	125	
Ser Arg Pro Ser Thr Ser Asp His Leu Arg Lys Ser Ser Glu Ile Val			
130	135	140	
Leu Met Ser Leu Gln Ile Ile Gly Leu Ser Phe Leu Arg Val Thr Lys			
145	150	155	160
Ile Asn Ile Glu Leu Thr Asn Arg Asp Val Thr Thr Leu Leu Leu Phe			
165	170	175	
Trp Leu Ile Leu Leu Ser Leu Ser Ile Leu Arg Val Tyr Lys Arg Ser			
180	185	190	
Thr Asn Leu Trp Ala Ile Cys Phe Thr Ala His Thr Thr Ile Trp Ile			
195	200	205	
Ser Thr Trp Ile Pro Ile Arg Ser Val Tyr Ile Gly Asn Ile Asp Asp			
210	215	220	
Val Pro Ser Gln Ile Phe Tyr Ile Phe Glu Phe Val Ile Thr Ser Thr			
225	230	235	240
Leu Gln Pro Ile Lys Leu Thr Ser Pro Ile Lys Asp Asn Ser Ser Ile			
245	250	255	
Ile Tyr Val Arg Asp Asp His Thr Ser Pro Ser Arg Glu His Ile Ser			
260	265	270	
Ser Ile Leu Ser Cys Ile Thr Trp Ser Trp Ile Thr Asn Phe Ile Trp			
275	280	285	

Glu Ala Gln Lys Asn Thr Ile Lys Leu Lys Asp Ile Trp Gly Leu Ser
 290 295 300

Met Glu Asp Tyr Ser Ile Phe Ile Leu Lys Gly Phe Thr Arg Arg Asn
 305 310 315 320

Lys His Ile Asn Asn Leu Thr Leu Ala Leu Phe Glu Ser Phe Lys Thr
 325 330 335

Tyr Leu Leu Ile Gly Met Leu Trp Val Leu Val Asn Ser Ile Val Asn
 340 345 350

Leu Leu Pro Thr Ile Leu Met Lys Arg Phe Leu Glu Ile Val Asp Asn
 355 360 365

Pro Asn Arg Ser Ser Ser Cys Met Asn Leu Ala Trp Leu Tyr Ile Ile
 370 375 380

Gly Met Phe Ile Cys Arg Leu Thr Leu Ala Ile Cys Asn Ser Gln Gly
 385 390 395 400

Gln Phe Val Ser Asp Lys Ile Cys Leu Arg Ile Arg Ala Ile Leu Ile
 405 410 415

Gly Glu Ile Tyr Ala Lys Gly Leu Arg Arg Arg Leu Phe Thr Ser Pro
 420 425 430

Lys Thr Ser Ser Asp Ser Asp Ser Ile Ser Ala Asn Leu Gly Thr Ile
 435 440 445

Ile Asn Leu Ile Ser Ile Asp Ser Phe Lys Val Ser Glu Leu Ala Asn
 450 455 460

Tyr Leu Tyr Val Thr Val Gln Ala Val Ile Met Ile Ile Val Val Val
 465 470 475 480

Gly Leu Leu Phe Asn Phe Leu Gly Val Ser Ala Phe Ala Gly Ile Ser
 485 490 495

Ile Ile Leu Val Met Phe Pro Leu Asn Phe Leu Leu Ala Asn Leu Leu
 500 505 510

Gly Lys Phe Gln Lys Gln Thr Leu Lys Cys Thr Asp Gln Arg Ile Ser
 515 520 525

Lys Leu Asn Glu Cys Leu Gln Asn Ile Arg Ile Val Lys Tyr Phe Ala
 530 535 540

Trp Glu Arg Asn Ile Ile Asn Glu Ile Lys Ser Ile Arg Gln Lys Glu
 545 550 555 560

Leu Arg Ser Leu Leu Lys Lys Ser Leu Val Trp Ser Val Thr Ser Phe
 565 570 575

Leu Trp Phe Val Thr Pro Thr Leu Val Thr Gly Val Thr Phe Ala Ile
 580 585 590

Cys Thr Phe Val Gln His Glu Asp Leu Asn Ala Pro Leu Ala Phe Thr
 595 600 605

Thr Leu Ser Leu Phe Thr Leu Leu Lys Thr Pro Leu Asp Gln Leu Ser
 610 615 620

Asn Met Leu Ser Phe Ile Asn Gln Ser Lys Val Ser Leu Lys Arg Ile
 625 630 635 640

Ser Asp Phe Leu Arg Met Asp Asp Thr Glu Lys Tyr Asn Gln Leu Thr
 645 650 655

Ile Ser Pro Asp Lys Asn Lys Ile Glu Phe Lys Asn Ala Thr Leu Thr
 660 665 670

Trp Asn Glu Asn Asp Ser Asp Met Asn Ala Phe Lys Leu Cys Gly Leu
 675 680 685

Asn Ile Lys Phe Gln Ile Gly Lys Leu Asn Leu Ile Leu Gly Ser Thr
 690 695 700

Gly Ser Gly Lys Ser Ala Leu Leu Leu Gly Leu Leu Gly Glu Leu Asn
 705 710 715 720

Leu Ile Ser Gly Ser Ile Ile Val Pro Ser Leu Glu Pro Lys His Asp
 725 730 735

Leu Ile Pro Asp Cys Glu Gly Leu Thr Asn Ser Phe Ala Tyr Cys Ser
 740 745 750

Gln Ser Ala Trp Leu Leu Asn Asp Thr Val Lys Asn Asn Ile Ile Phe
 755 760 765

Asp Asn Phe Tyr Asn Glu Asp Arg Tyr Asn Lys Val Ile Asp Ala Cys
 770 775 780

Gly Leu Lys Arg Asp Leu Glu Ile Leu Pro Ala Gly Asp Leu Thr Glu
 785 790 795 800

Ile Gly Glu Lys Gly Ile Thr Leu Ser Gly Gly Gln Lys Gln Arg Ile
805 810 815

Ser Leu Ala Arg Ala Val Tyr Ser Ser Ala Lys His Val Leu Leu Asp
820 825 830

Asp Cys Leu Ser Ala Val Asp Ser His Thr Ala Val Trp Ile Tyr Glu
835 840 845

Asn Cys Ile Thr Gly Pro Leu Met Lys Asn Arg Thr Cys Ile Leu Val
850 855 860

Thr His Asn Val Ser Leu Thr Leu Arg Asn Ala His Phe Ala Ile Val
865 870 875 880

Leu Glu Asn Gly Lys Val Lys Asn Gln Gly Thr Ile Thr Glu Leu Gln
885 890 895

Ser Lys Gly Leu Phe Lys Glu Lys Tyr Val Gln Leu Ser Ser Arg Asp
900 905 910

Ser Ile Asn Glu Lys Asn Ala Asn Arg Leu Lys Ala Pro Arg Lys Asn
915 920 925

Asp Ser Gln Lys Ile Glu Pro Val Thr Glu Asn Ile Asn Phe Asp Ala
930 935 940

Asn Phe Val Asn Asp Gly Gln Leu Ile Glu Glu Glu Glu Lys Ser Asn
945 950 955 960

Gly Ala Ile Ser Pro Asp Val Tyr Lys Trp Tyr Leu Lys Phe Phe Gly
965 970 975

Gly Phe Lys Ala Leu Thr Ala Leu Phe Ala Leu Tyr Ile Thr Ala Gln
980 985 990

Ile Leu Phe Ile Ser Gln Ser Trp Trp Ile Arg His Trp Val Asn Asp
995 1000 1005

Thr Asn Val Arg Ile Asn Ala Pro Gly Phe Ala Met Asp Thr Leu
1010 1015 1020

Pro Leu Lys Gly Met Thr Asp Ser Ser Lys Asn Lys His Asn Ala
1025 1030 1035

Phe Tyr Tyr Leu Thr Val Tyr Phe Leu Ile Gly Ile Ile Gln Ala

1040		1045		1050
Met Leu Gly Gly Phe Lys Thr	Met Met Thr Phe Leu	Ser Gly Met		
1055	1060	1065		
Arg Ala Ser Arg Lys Ile Phe	Asn Asn Leu Leu Asp	Leu Val Leu		
1070	1075	1080		
His Ala Gln Ile Arg Phe Phe	Asp Val Thr Pro Val	Gly Arg Ile		
1085	1090	1095		
Met Asn Arg Phe Ser Lys Asp	Ile Glu Gly Val Asp	Gln Glu Leu		
1100	1105	1110		
Ile Pro Tyr Leu Glu Val Thr	Ile Phe Cys Leu Ile	Gln Cys Ala		
1115	1120	1125		
Ser Ile Ile Phe Leu Ile Thr	Val Ile Thr Pro Arg	Phe Leu Thr		
1130	1135	1140		
Val Ala Val Ile Val Phe Val	Leu Tyr Phe Phe Val	Gly Lys Trp		
1145	1150	1155		
Tyr Leu Thr Ala Ser Arg Glu	Leu Lys Arg Leu Asp	Ser Ile Thr		
1160	1165	1170		
Lys Ser Pro Ile Phe Gln His	Phe Ser Glu Thr Leu	Val Gly Val		
1175	1180	1185		
Cys Thr Ile Arg Ala Phe Gly	Asp Glu Arg Arg Phe	Ile Leu Glu		
1190	1195	1200		
Asn Met Asn Lys Ile Asp Gln	Asn Asn Arg Ala Phe	Phe Tyr Leu		
1205	1210	1215		
Ser Val Thr Val Lys Trp Phe	Ser Phe Arg Val Asp	Met Ile Gly		
1220	1225	1230		
Ala Phe Ile Val Leu Ala Ser	Gly Ser Phe Ile Leu	Leu Asn Ile		
1235	1240	1245		
Ala Asn Ile Asp Ser Gly Leu	Ala Gly Ile Ser Leu	Thr Tyr Ala		
1250	1255	1260		
Ile Leu Phe Thr Asp Gly Ala	Leu Trp Leu Val Arg	Leu Tyr Ser		
1265	1270	1275		

Thr	Phe	Glu	Met	Asn	Met	Asn	Ser	Val	Glu	Arg	Leu	Lys	Glu	Tyr
1280						1285					1290			
Ser	Ser	Ile	Glu	Gln	Glu	Asn	Tyr	Leu	Gly	His	Asp	Glu	Gly	Arg
1295						1300					1305			
Ile	Leu	Leu	Leu	Asn	Glu	Pro	Ser	Trp	Pro	Lys	Asp	Gly	Glu	Ile
1310						1315					1320			
Glu	Ile	Glu	Asn	Leu	Ser	Leu	Arg	Tyr	Ala	Pro	Asn	Leu	Pro	Pro
1325						1330					1335			
Val	Ile	Arg	Asn	Val	Ser	Phe	Lys	Val	Asp	Pro	Gln	Ser	Lys	Ile
1340						1345					1350			
Gly	Ile	Val	Gly	Arg	Thr	Gly	Ala	Gly	Lys	Ser	Thr	Ile	Ile	Thr
1355						1360					1365			
Ala	Leu	Phe	Arg	Leu	Leu	Glu	Pro	Ile	Thr	Gly	Cys	Ile	Lys	Ile
1370						1375					1380			
Asp	Gly	Gln	Asp	Ile	Ser	Lys	Ile	Asp	Leu	Val	Thr	Leu	Arg	Arg
1385						1390					1395			
Ser	Ile	Thr	Ile	Ile	Pro	Gln	Asp	Pro	Ile	Leu	Phe	Ala	Gly	Thr
1400						1405					1410			
Ile	Lys	Ser	Asn	Val	Asp	Pro	Tyr	Asp	Glu	Tyr	Asp	Glu	Lys	Lys
1415						1420					1425			
Ile	Phe	Lys	Ala	Leu	Ser	Gln	Val	Asn	Leu	Ile	Ser	Ser	His	Glu
1430						1435					1440			
Phe	Glu	Glu	Val	Leu	Asn	Ser	Glu	Glu	Arg	Phe	Asn	Ser	Thr	His
1445						1450					1455			
Asn	Lys	Phe	Leu	Asn	Leu	His	Thr	Glu	Ile	Ala	Glu	Gly	Gly	Leu
1460						1465					1470			
Asn	Leu	Ser	Gln	Gly	Glu	Arg	Gln	Leu	Leu	Phe	Ile	Ala	Arg	Ser
1475						1480					1485			
Leu	Leu	Arg	Glu	Pro	Lys	Ile	Ile	Leu	Leu	Asp	Glu	Ala	Thr	Ser
1490						1495					1500			
Ser	Ile	Asp	Tyr	Asp	Ser	Asp	His	Leu	Ile	Gln	Gly	Ile	Ile	Arg
1505						1510					1515			

Ser Glu Phe Asn Lys Ser Thr Ile Leu Thr Ile Ala His Arg Leu
1520 1525 1530

Arg Ser Val Ile Asp Tyr Asp Arg Ile Ile Val Met Asp Ala Gly
1535 1540 1545

Glu Val Lys Glu Tyr Asp Arg Pro Ser Glu Leu Leu Lys Asp Glu
1550 1555 1560

Arg Gly Ile Phe Tyr Ser Met Cys Arg Asp Ser Gly Gly Leu Glu
1565 1570 1575

Leu Leu Lys Gln Ile Ala Lys Gln Ser Ser Lys Met Met Lys
1580 1585 1590

<210> 84

<211> 2000

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 84

```

atgggaacgg atcccccttat tatccgaaat aatgggtcat tttgggaagt tgatgatttt      60
actcgtttag gaagaactca gctattgagc tactattttac cattggctat catagcctca      120
attggcattt tcgcactttg tcgcagtgga ttatctcggt atgtaagatc tgccgagtgc      180
gatttagtga acgaatatct atttggcgca caagaagaga gaaaagaaga taatagtata      240
gaaagacttc tacggaactc aaatacccaa gccaatcacg tcaacgtcaa aaagcaagga      300
aggattttga aacttagaca ttttgatata acaactatag atgtcaagca aatcgatgct      360
aaaaatcatg gtggactaac gtttagtaga ccgctacta gtgaccactt aagaaaatca      420
tctgaaattg tattaatgtc tttaaaaata attggccttt cctttttaag agtaacaaaa      480
atcaatattg aattaacgaa cagagatggt acaactttac tattattttg gttaatacta      540
ctttccctaa gtatcttaag agtttacaag cgttcaacga atctttgggc catctgtttt      600
actgcccata caactatttg gatttcaacc tggattccaa ttcgttcggt ctatattggt      660
aatatcgatg atgtaccctc acagatatct tacatctttg aattcgtaat tacttcaacc      720
ttacagccaa taaagctcac ttcaccgatt aaagacaact catccatcat ctacgttaga      780
gacgaccata cgtctccttc gagggaaacac atatcctcaa ttttaagttg cattacttgg      840
agctggatta ccaattttat atgggaggcc caaaagaaca ctattaagtt aaaggatatt      900
tggggcttat caatggaaga ctatagcatt ttcattctaa aagggtttac caggagaaac      960
aagcacatta ataatttgac gctagcactg tttgaatctt tcaaaacata tttactcata     1020
ggaatgttat gggttctggt gaacagtatt gtaaaccctc ttcccacaat tttaatgaaa     1080

```

```

agatttttag aaattgtgga taacccaaac cgttcctcat catgcatgaa tttggcgtgg 1140
ctttatatta ttggtatggt catttgtaga ttgacattag caatttgtaa ttcccaaggt 1200
caatttgttt ctgataagat ttgtttaaga ataagagcca tactcatagg agaaatttat 1260
gcaaaaggct tacgtaggag gctgtttaca tctccaaaaa ccagctctga ttcagatagt 1320
atctccgcaa accttggtac cataattaat ctcatattcta ttgactcatt taaggatatcg 1380
gaactagcaa actaccttta tgtgacagtt caggcagtaa ttatgataat agttgttgta 1440
ggactacttt tcaacttttt aggtgtttca gcttttgcag gaatttcaat tatcttagtg 1500
atgttcccat tgaatttctt gttagcgaat ttgttaggta agtttcaaaa gcaaactg 1560
aaatgtactg accaaagaat ctcaaaattg aacgagtgtc tacagaacat aagaattgtc 1620
aaatattttg cttgggaaag gaatattata aatgaaatca aatcaataag gcaaaaggaa 1680
ttaagatcct tattaaaaaa atctttggtg tgggtccgtaa cttcttttct ttggttcgtg 1740
acaccgacct tggtgacagg tgtcactttc gccatctgta catttggtca acatgaagat 1800
ttgaatgccc cgcttgcttt cactactttg tcactcttca ctttggttaa gacaccctg 1860
gatcaattat caaatatgct aagtttcata aatcaatcaa aagtctctct aaaaagaata 1920
agcgattttt taaggatgga cgatacagaa aaatataatc aactaaccat atctccagac 1980
aaaaataaaa ttgaatttaa 2000

```

<210> 85

<211> 329

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 85

```

Met Thr Thr Val Ser Ile Asn Lys Pro Asn Leu Leu Lys Phe Lys His
1           5           10           15

```

```

Val Lys Ser Phe Gln Pro Gln Glu Lys Asp Cys Gly Pro Val Thr Ser
20           25           30

```

```

Leu Asn Phe Asp Asp Asn Gly Gln Phe Leu Leu Thr Ser Ser Ser Asn
35           40           45

```

```

Asp Thr Met Gln Leu Tyr Ser Ala Thr Asn Cys Lys Phe Leu Asp Thr
50           55           60

```

```

Ile Ala Ser Lys Lys Tyr Gly Cys His Ser Ala Ile Phe Thr His Ala
65           70           75           80

```

```

Gln Asn Glu Cys Ile Tyr Ser Ser Thr Met Lys Asn Phe Asp Ile Lys
85           90           95

```

Tyr Leu Asn Leu Glu Thr Asn Gln Tyr Leu Arg Tyr Phe Ser Gly His
 100 105 110

Gly Ala Leu Val Asn Asp Leu Lys Met Asn Pro Val Asn Asp Thr Phe
 115 120 125

Leu Ser Ser Ser Tyr Asp Glu Ser Val Arg Leu Trp Asp Leu Lys Ile
 130 135 140

Ser Lys Pro Gln Val Ile Ile Pro Ser Leu Val Pro Asn Cys Ile Ala
 145 150 155 160

Tyr Asp Pro Ser Gly Leu Val Phe Ala Leu Gly Asn Pro Glu Asn Phe
 165 170 175

Glu Ile Gly Leu Tyr Asn Leu Lys Lys Ile Gln Glu Gly Pro Phe Leu
 180 185 190

Ile Ile Lys Ile Asn Asp Ala Thr Phe Ser Gln Trp Asn Lys Leu Glu
 195 200 205

Phe Ser Asn Asn Gly Lys Tyr Leu Leu Val Gly Ser Ser Ile Gly Lys
 210 215 220

His Leu Ile Phe Asp Ala Phe Thr Gly Gln Gln Leu Phe Glu Leu Ile
 225 230 235 240

Gly Thr Arg Ala Phe Pro Met Arg Glu Phe Leu Asp Ser Gly Ser Ala
 245 250 255

Cys Phe Thr Pro Asp Gly Glu Phe Val Leu Gly Thr Asp Tyr Asp Gly
 260 265 270

Arg Ile Ala Ile Trp Asn His Ser Asp Ser Ile Ser Asn Lys Val Leu
 275 280 285

Arg Pro Gln Gly Phe Ile Pro Cys Val Ser His Glu Thr Cys Pro Arg
 290 295 300

Ser Ile Ala Phe Asn Pro Lys Tyr Ser Met Phe Val Thr Ala Asp Glu
 305 310 315 320

Thr Val Asp Phe Tyr Val Tyr Asp Glu
 325

<210> 86
 <211> 990

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 86

```

atgaccaccg tgtccatcaa taagcccaac ctgctgaaat tcaagcatgt taaaagcttt      60
caacctcaag aaaaagactg cggacccgta acctcattga atttcgacga taatggccag      120
tttctactga cctcttcttc caacgatata atgcaattgt acagtgccac gaactgcaaa      180
ttcttggaca ctatagcctc taagaaatat ggctgtcact ccgtatctt tacgcacgca      240
caaaacgaat gtatctattc ctctacaatg aaaaattttg acattaaata ccttaatctg      300
gaaacaaacc aatatctaag atatttttcc ggtcatggcg ccctagttaa tgatttgaag      360
atgaaccccg tgaacgatac gtttctatcg tcgtcatacg atgaatccgt taggctttgg      420
gatttgaaga tctctaaacc gcaagttatt ataccaagtc tcgtacaaa ttgtatcgca      480
tatgatccaa gtggccttgt attcgattg gggaaccag agaatttcga aatagggcta      540
tataatctga aaaaaattca ggagggtcct ttcttgataa ttaaaattaa tgatgcgact      600
ttcagtcaat ggaataaatt agaattttct aacaatggaa agtatttatt agttggctcc      660
tcgataggaa agcatttaat ttttgacgca ttcacaggtc aacaattatt cgaactaata      720
ggaacaaggg ccttcccgat gagagaattt ctagattctg gatctgcttg tttcacacca      780
gatggtgaat tcgtccttgg aaccgattat gacggtagga ttgccatttg gaatcattct      840
gattcaataa gtaacaaagt attaaggccg caagggttca ttccctgtgt ttctcatgag      900
acctgcccc aagtcaattgc attcaacctt aaatattcga tgtttgttac cgcagacgaa      960
acagtagatt ttacgtgta cgatgaatga                                     990

```

<210> 87

<211> 220

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 87

```

Met Ser Ala Gly Asp Ile Ser Ala Ile Asn Ile Lys Ser Val Lys Lys
1           5           10           15

Asn Arg Arg Arg Lys Lys Arg Arg Thr Ala Asp Val Ser Ser Ser Asp
          20           25           30

Ser Ser Ser Ser Asp Pro Ser Ser Glu Ser Glu Lys Glu Glu Ile Gln
          35           40           45

Asn Gly Ala Ile Glu Glu His Val Gly Glu Asn Gly Lys Ser Asp His
          50           55           60

Val Phe Ser Lys Gly Asn Asp Glu Asp Lys Gln Glu Asp Ile Ala Ile

```

65		70		75		80
Glu Val Ser Asp	Val Glu Leu Thr Asp	Glu Glu Ser Lys Asp	Leu Lys			
	85	90	95			
Leu Asn Ser Lys	Glu Val Ile Asp Asp	Leu Thr Lys Ile	Ser Leu Ser			
	100	105	110			
Lys Ile Pro Glu	Pro Thr Lys Ser Gln Asn	Lys Glu Gly Phe Met Asn				
	115	120	125			
Ala Ser Lys Ile	Ala Glu Asn Ile Lys Leu Ala Arg	Glu Glu Tyr Asn				
	130	135	140			
Glu Leu Ala Glu	Asn Phe Val Pro Lys Gly Lys Asp	Lys Thr Lys Leu				
145	150	155	160			
Arg Glu Glu Tyr	Leu Asn Leu Leu Phe Glu Asn Tyr Gly Asp	Asp Ile				
	165	170	175			
Asn Arg Leu Arg	Ala Ala Pro Asp Phe Thr Asn Lys Ser	Leu Ser Ile				
	180	185	190			
Leu Ala Asp Ala	Leu Gln Glu Gly Ile Gly Met Phe Asp	Ile Gly Glu				
	195	200	205			
Leu Glu Leu Val	Leu Lys Asn Lys Glu Met Glu Asn					
	210	215	220			

<210> 88

<211> 663

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 88

atgtcggcag gtgatataat agccataaat atcaagtctg tcaaaaaaaaa cagaaggagg	60
aagaagagaa gaacagctga tgtttcatca tcagattctt catcatcgga tccatcatca	120
gaaagtgaaa aggaggaaat ccaaaatggg gccatcgaag aacacgttgg agaaaatggt	180
aaaagtgatc atgttttctc aaaaggtaat gacgaagaca aacaagagga cattgcaata	240
gaagtttcgg atgtcgagct tacagacgaa gaaagcaagg atttgaagtt aaattcaaaa	300
gaagtgatag acgatttaac caaaatttct ttgagcaaga tcccagagcc tacgaaatct	360
caaaacaagg agggttttat gaatgcatcg aaaattgccg aaaatatcaa gcttgcgaga	420
gaagaataca atgaattggc agaaaacttt gtgcccagg ggaaagacaa gacaaagtta	480
agggagaagt acttaaattt actttttgaa aactacggtg atgatatcaa tcgtcttaga	540

143/148

gctgccccgg atttcacgaa taaatcacta tccattttgg cagatgctct gcaggaaggc 600
 ataggaatgt ttgatattgg tgaactagaa ttggtcttga aaaataaaga aatggagaac 660
 tga 663

<210> 89
 <211> 441
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 89

Met Ser Ser Thr Ile Phe Tyr Arg Phe Lys Ser Gln Arg Asn Thr Ser
 1 5 10 15

Arg Ile Leu Phe Asp Gly Thr Gly Leu Thr Val Phe Asp Leu Lys Arg
 20 25 30

Glu Ile Ile Gln Glu Asn Lys Leu Gly Asp Gly Thr Asp Phe Gln Leu
 35 40 45

Lys Ile Tyr Asn Pro Asp Thr Glu Glu Glu Tyr Asp Asp Ala Phe
 50 55 60

Val Ile Pro Arg Ser Thr Ser Val Ile Val Lys Arg Ser Pro Ala Ile
 65 70 75 80

Lys Ser Phe Ser Val His Ser Arg Leu Lys Gly Asn Val Gly Ala Ala
 85 90 95

Ala Leu Gly Asn Ala Thr Arg Tyr Val Thr Gly Arg Pro Arg Val Leu
 100 105 110

Gln Lys Arg Gln His Thr Ala Thr Thr Thr Ala Asn Val Ser Gly Thr
 115 120 125

Thr Glu Glu Glu Arg Ile Ala Ser Met Phe Ala Thr Gln Glu Asn Gln
 130 135 140

Trp Glu Gln Thr Gln Glu Glu Met Ser Ala Ala Thr Pro Val Phe Phe
 145 150 155 160

Lys Ser Gln Thr Asn Lys Asn Ser Ala Gln Glu Asn Glu Gly Pro Pro
 165 170 175

Pro Pro Gly Tyr Met Cys Tyr Arg Cys Gly Gly Arg Asp His Trp Ile
 180 185 190

Lys Asn Cys Pro Thr Asn Ser Asp Pro Asn Phe Glu Gly Lys Arg Ile

195	200	205
Arg Arg Thr Thr Gly Ile Pro Lys Lys Phe Leu Lys Ser Ile Glu Ile		
210	215	220
Asp Pro Glu Thr Met Thr Pro Glu Glu Met Ala Gln Arg Lys Ile Met		
225	230	235 240
Ile Thr Asp Glu Gly Lys Phe Val Val Gln Val Glu Asp Lys Gln Ser		
	245	250 255
Trp Glu Asp Tyr Gln Arg Lys Arg Glu Asn Arg Gln Ile Asp Gly Asp		
	260	265 270
Glu Thr Ile Trp Arg Lys Gly His Phe Lys Asp Leu Pro Asp Asp Leu		
	275	280 285
Lys Cys Pro Leu Thr Gly Gly Leu Leu Arg Gln Pro Val Lys Thr Ser		
	290	295 300
Lys Cys Cys Asn Ile Asp Phe Ser Lys Glu Ala Leu Glu Asn Ala Leu		
	305	310 315 320
Val Glu Ser Asp Phe Val Cys Pro Asn Cys Glu Thr Arg Asp Ile Leu		
	325	330 335
Leu Asp Ser Leu Val Pro Asp Gln Asp Lys Glu Lys Glu Val Glu Thr		
	340	345 350
Phe Leu Lys Lys Gln Glu Glu Leu His Gly Ser Ser Lys Asp Gly Asn		
	355	360 365
Gln Pro Glu Thr Lys Lys Met Lys Leu Met Asp Pro Thr Gly Thr Ala		
	370	375 380
Gly Leu Asn Asn Asn Thr Ser Leu Pro Thr Ser Val Asn Asn Gly Gly		
	385	390 395 400
Thr Pro Val Pro Pro Val Pro Leu Pro Phe Gly Ile Pro Pro Phe Pro		
	405	410 415
Met Phe Pro Met Pro Phe Met Pro Pro Thr Ala Thr Ile Thr Asn Pro		
	420	425 430
His Gln Ala Asp Ala Ser Pro Lys Lys		
	435	440

<210> 90
 <211> 1326
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 90
 atgagtagca cgatatttta cgcctttaag tctcaacgaa acacatcaag aattttattt 60
 gatggtaccg gcctgacagt atttgatttg aaaagggaaa ttattcaaga gaacaaacta 120
 ggtgacggca cagattttcca attaaaaatt tacaaccagc atacagaaga ggaatacagc 180
 gatgatgcct ttgttatacc tagatctact agtgtcatag taaaagatc tccagcaatt 240
 aaatcattct ccgtacacag tgcacttaa gggaaatgtgg gagcagcagc tcttgggaac 300
 gcaacaaggt atgttactgg taggccaaga gtgttgcaaa agagacaaca cactgctaca 360
 accactgcta atgttagtgg tacaacggaa gaagaaagaa ttgctagtat gtttgccaca 420
 caagaaaatc aatgggaaca aacgcaagaa gaaatgtctg cagccacacc tgtttttttc 480
 aagtcacaga cgaataagaa ttctgcacaa gaaaacgaag gccaccgcc accaggttat 540
 atgtgctatc gttgtggggg tagagaccac tggattaaaa attgtccaac taacagcgat 600
 ccaaattttg aaggaaaaag aatcagaaga accacaggta ttccaaagaa gtttttaaaa 660
 tccattgaaa tagatccga gacaatgaca ccggaagaga tggctcagcg aaagattatg 720
 attacggacg aaggcaagtt cgtggtacaa gttgaagaca aacaatcatg ggaagactac 780
 caaaggaaaa gagagaaccg tcaaattgat ggtgatgaaa ccatttggag aaaaggccat 840
 ttcaaagatc ttctgacga tttaaaatgt cccttgacag gtggtctttt gaggcagccg 900
 gtaaagacaa gcaagtgtctg taacatagat ttctcaaaag aggcgctgga aatgcactg 960
 gttagagagcg actttgtatg cccaattgc gaaaccgcg atatccttct cgattcttta 1020
 gtaccgcgacc aggacaagga aaaggaggtc gaaacgtttt tgaagaaaca agaggaacta 1080
 cacggaagct ctaaagatgg caaccagcca gaaactaaga aatgaagtt gatggatcca 1140
 actggcaccg ctggcttgaa caacaatacc agccttccaa cttctgtaaa taacggcgggt 1200
 acgccagtgc caccagtacc gttaccttcc ggtatacctc ctttcccat gtttccaatg 1260
 cccttcatgc ctccaacggc tactatcaca aatcctcacc aagctgacgc aagccctaag 1320
 aaatga 1326

<210> 91
 <211> 159
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 91

Met Ser Arg Met Pro Ser Ser Phe Asp Val Thr Glu Arg Asp Leu Asp
 1 5 10 15

Asp Met Thr Phe Gly Glu Arg Ile Ile Tyr His Cys Lys Lys Gln Pro
 20 25 30

Leu Val Pro Ile Gly Cys Leu Leu Thr Thr Gly Ala Val Ile Leu Ala
 35 40 45

Ala Gln Asn Val Arg Leu Gly Asn Lys Trp Lys Ala Gln Tyr Tyr Phe
 50 55 60

Arg Trp Arg Val Gly Leu Gln Ala Ala Thr Leu Val Ala Leu Val Ala
 65 70 75 80

Gly Ser Phe Ile Tyr Gly Thr Ser Gly Lys Glu Leu Lys Ala Lys Glu
 85 90 95

Glu Gln Leu Lys Glu Lys Ala Lys Met Arg Glu Lys Leu Trp Ile Gln
 100 105 110

Glu Leu Glu Arg Arg Glu Glu Glu Thr Glu Ala Arg Arg Lys Arg Ala
 115 120 125

Glu Leu Ala Arg Met Lys Thr Leu Glu Asn Glu Glu Glu Ile Lys Asn
 130 135 140

Leu Glu Lys Glu Leu Ser Asp Leu Glu Asn Lys Leu Gly Lys Lys
 145 150 155

<210> 92
 <211> 480
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 92
 atgtcacgca tgccatctag ttctgatgtt acggagaggg atttggatga tatgaccttt 60
 ggcgaaagga ttatatacca ttgtaagaaa cagccattgg taccattgg gtgcttgctg 120
 actacaggag ctgtcattct ggctgctcaa aatgttcgtc ttggaataa atggaaagct 180
 cagtactact tccgttggcg tgtgggtcta caagcggcca cactagtcgc actagtcgca 240
 ggttcattta tctatgggac ttctggtaag gaactgaagg cgaaggagga acaattgaag 300
 gagaaagcca agatgagaga aaagttatgg atccaagagc tggagagaag ggaggaagaa 360
 acggaggcaa ggagaaaaag agccgaattg gcaagaatga agacccttga gaacgaagag 420
 gaaatcaaga acttagaaaa ggaactaagc gacctggaaa ataagcttgg aaagaagtaa 480

<210> 93
 <211> 188
 <212> PRT

<213> *Saccharomyces cerevisiae*

<400> 93

Met Asp Leu Pro Lys Asp Lys Ser Asp Arg Thr His Gln Arg Ile Asn
 1 5 10 15

Leu Asn Asn Ser Gly Thr Asp Arg Thr Asn Asp Leu Tyr Leu His Ile
 20 25 30

Val Gln Thr Phe Gly Cys Ile Glu Thr Thr Ala Thr Glu Asn Ala Thr
 35 40 45

Lys Leu Leu Met Leu Gly Asp Val Glu Val Glu Ile Ser Ala Ser Ser
 50 55 60

Val Ser Ile Glu Trp Thr Gln Lys Ser Met Ile Ser Gln Thr Ile Ala
 65 70 75 80

Asp Ser Ile Val Ile Met Ile Ile Gly Leu Cys Ala Ser Asp Lys Asn
 85 90 95

Val Leu Ser Glu Ser Glu Leu Lys Glu Arg Asn His Asn Val Trp Lys
 100 105 110

Ile Gln Glu Leu Gln Asn Leu Phe Arg Glu Gln Phe Gly Asp Ser Phe
 115 120 125

Ser Ile Asp Glu Gly Ile Gly Lys Lys Glu Asn Val Lys Asn Gly Ser
 130 135 140

Val Thr Ile Gly Lys Ser Lys Ala Thr Ile Asp Phe Ser Thr Met Lys
 145 150 155 160

Leu Ile Asp Cys Asn Ser Asn Pro Leu Lys Gly Arg Val Glu Ser Ile
 165 170 175

Leu Ser Ile Gly Gln Lys Leu Thr Thr Pro Leu Cys
 180 185

<210> 94

<211> 567

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 94

atggatttac ccaaggacaa aagtacaga actcatcaaa gaattaatct aaataacagt 60

gggacagatc gaactaatga ttgtacctt catattgtcc aaacgttcgg ttgcatagaa 120

acaactgcaa cggaaaatgc caccgaaactg ttaatgctgg gtgacgtcga agtagaaata 180

tctgcgagca gcgtttcaat tgagtggaca cagaagtcaa tgataagcca aacaattgcc	240
gatagtatag taataatgat catcggtttg tgtgcaagcg acaagaacgt gctatctgaa	300
tcagaattga aagagagaaa ccataacggtt tggaagatcc aagaattgca aaatctgttt	360
cgagaacaat ttggagacag ttttagcatc gatgaaggaa taggaaaaaa agaaaatgta	420
aagaatggta gcgtcaccat aggcaagagt aaagccacga tcgatttctc caccatgaag	480
ctgattgatt gtaattcgaa ccactaaag ggaagagtgg agagcatact aagcattggc	540
cagaaattaa caactccatt gtgctga	567

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 November 2002 (21.11.2002)

PCT

(10) International Publication Number
WO 02/092626 A3

(51) International Patent Classification⁷: **C07K 14/39**,
14/47, C12N 15/11, 15/62

(74) Agent: **GRÜNECKER, KINKELDEY, STOCKMAIR
& SCHWANHÄUSSER**; Maximilianstrasse 58, 80538
München (DE).

(21) International Application Number: **PCT/EP02/05359**

(22) International Filing Date: **15 May 2002 (15.05.2002)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:
01111774.4 **15 May 2001 (15.05.2001)** EP
01130253.6 **20 December 2001 (20.12.2001)** EP

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **CELL-ZOME AG** [DE/DE]; Meyerhofstrasse 1, 69117 Heidelberg (DE).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **BAUER, Andreas** [DE/DE]; Dammweg 9, 69123 Heidelberg (DE). **GAVIN, Anne-Claude** [CH/DE]; Heltenstrasse 44, 69181 Leimen (DE). **MARZIOCH, Martina** [DE/DE]; Berghalde 34, 69126 Heidelberg (DE). **SCHULTZ, Jörg** [DE/DE]; Im Beind 21, 69126 Heidelberg (DE). **BRAJENOVIC, Miro** [HR/DE]; Am Rosenbusch 4, 69117 Heidelberg (DE). **GRANDI, Paola** [IT/DE]; Klingenteichstrasse 22, 69117 Heidelberg (DE).

Published:

— *with international search report*

(88) Date of publication of the international search report:
27 November 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/092626 A3

(54) Title: **CLEAVAGE AND POLYADENYLATION COMPLEX OF PRECURSOR MRNA**

(57) Abstract: The present invention relates to novel components of the cleavage/polyadenylation machinery of precursor mRNA as well as to the complex containing the new components and its use. The complex is obtained by using one component thereof as a bait and isolating a highly organised complex consisting of at least 13 distinct proteins.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/05359

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/39 C07K14/47 C12N15/11 C12N15/62		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C07K C12N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, BIOSIS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	VO LE THUY ANH ET AL: "Mpe1, a zinc knuckle protein, is an essential component of yeast cleavage and polyadenylation factor required for the cleavage and polyadenylation of mRNA." MOLECULAR AND CELLULAR BIOLOGY, vol. 21, no. 24, December 2001 (2001-12), pages 8346-8356, XP002223058 ISSN: 0270-7306	1-10, 12-19, 30-34, 36-38
Y	Abstract, Introduction, Methods Fig.2-7 and legends, Discussion. --- -/--	40-45
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *Z* document member of the same patent family		
Date of the actual completion of the international search 9 December 2002		Date of mailing of the international search report 17/01/2003
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Bretherick, J

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/05359

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE GENBANK 'Online! NCBI; Ygr156w, 11 August 1997 (1997-08-11) DYCK ET AL.: "Hypothetical Protein 'Saccharomyces cerevisiae!" retrieved from NCBI Database accession no. CAA97170 XP002223846 cited in the application URL www.ncbi.nlm.nih.gov:80/ See sequence listing</p> <p>---</p>	41-43
X	<p>DATABASE GENBANK 'Online! NCBI; Ylr221cp, 2 February 2001 (2001-02-02) JOHNSTON ET AL.: "Yr221cp 'saccharomyces cerevisiae!" retrieved from NCBI Database accession no. AAB67410 XP002223848 cited in the application URL www.ncbi.nlm.nih.gov:80/ See sequence listing</p> <p>---</p>	41-43
X	<p>DATABASE GENBANK 'Online! NCBI; YKL018w, 11 August 1997 (1997-08-11) RIEGER, M.: "ORF YKL018w 'Saccharomyces cerevisiae!" retrieved from NCBI Database accession no. CAA81853 XP002223849 cited in the application URL www.ncbi.nlm.nih.gov:80/ see sequence listing</p> <p>---</p>	41-43
Y	<p>ZHAO JING ET AL: "Formation of mRNA 3' ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis" MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 63, no. 2, June 1999 (1999-06), pages 405-445, XP002179541 ISSN: 1092-2172 See whole document, esp. pp.413-417, Table 3, Fig 4 and legend.</p> <p>---</p> <p style="text-align: center;">-/--</p>	1-10, 12-19, 30-34, 36-38, 40-45

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/05359

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>RIGAUT G ET AL: "A generic protein purification method for protein complex characterization and proteome exploration" NATURE BIOTECHNOLOGY, NATURE PUBLISHING, US, vol. 17, no. 10, October 1999 (1999-10), pages 1030-1032, XP002179540 ISSN: 1087-0156 cited in the application The whole document, especially Introduction and Fig 1 and legend.</p> <p>---</p>	1-10, 12-19, 30-34, 36-38, 40-45
Y	<p>KESSLER MARCO M ET AL: "Purification of the Saccharomyces cerevisiae cleavage/polyadenylation factor I: Separation into two components that are required for both cleavage and polyadenylation of mRNA 3' ends." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 271, no. 43, 1996, pages 27167-27175, XP002223060 ISSN: 0021-9258 Abstract, Fig 1 and legend, results, Figs. 1-8 and legends. page 0</p> <p>---</p>	1-10, 12-19, 30-34, 36-38, 40-45
Y	<p>ZHAO, J. ET AL.: "Pta1, a Component of Yeast CF II, Is Required for Both Cleavage and Poly(A) Addition of mRNA Precursor." MOLECULAR AND CELLULAR BIOLOGY, vol. 19, no. 11, November 1999 (1999-11), pages 7733-7740, XP002223845 Abstract, Results, esp. Fig.1,2,8 and legends.</p> <p>---</p>	1-10, 12-19, 30-34, 36-38, 40-45
Y	<p>TOLLERVEY D ET AL: "RNA PROCESSING MARCHES ON" CELL, CELL PRESS, CAMBRIDGE, MA, US, vol. 103, no. 5, 22 November 2000 (2000-11-22), pages 703-709, XP001021768 ISSN: 0092-8674 Last paragraph page 705 to penultimate paragraph on page 707, Fig. 2 and legend.</p> <p>-----</p>	1-10, 12-19, 30-34, 36-38, 40-45

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 02/05359

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 20-28
because they relate to subject matter not required to be searched by this Authority, namely:
Rule 39.1(iv) PCT - Diagnostic method practised on the human or animal body (claims 20-25) Rule 39.1(iv) PCT - Method for treatment of the human or animal body by therapy (claims 26-28)
2. ☒ Claims Nos.: 11, 20-29, 35, 39
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 11,20-29,35,39

Present claims 1-10,12-19, 30-34 and 36-38 and 40 are directed to an extremely large number of possible compounds and/or methods/uses. In fact, the claims contain so many options, undefined mammalian homologs, components of the complex considered similar solely by virtue of their encoding DNA hybridising under low stringency conditions and possible permutations that a lack of clarity (and/or conciseness) within the meaning of Article 6 PCT arises to such an extent as to render a meaningful search of the claims impossible. Consequently, the search has been carried out for those parts of the above claims which do appear to be clear and/or concise and or which find support in the description, namely the complexe(s) indicated in the examples. The same applies, mutatis mutandis, to any dependent claims.

Claim 39, is directed to a generic component of the polyadenylation-complex obtainable by the generic processes according to claims 36-38. Such a product is defined only in these broad terms and has therefore not been searched.

Claims 11, 20-29 and 35 have not been subject to a search. Claims 20-25 are directed to methods of diagnosis of disease. There are no direct clear and precise teachings pertaining to the use of the complexes and/or derivatives thereof in methods of diagnosis. The same applies to the subject-matter of claims 11, 26-29 and 35, which is directed to methods of treatment on the human and/or animal body and/or pharmaceutical compositions. There is no clear teaching pertaining to precise indications of disease and modi of therapy, nor to pharmaceutical compositions. The skilled person would be unable to ascertain such without either undue burden of experimentation or the use of inventive skill. Therefore, in the effective absence of such teaching, it is neither feasible nor worthwhile to carry out a search for such methods and/or compositions.

Claim 41 refers in a) to SEQ ID numbers which are easily identified as individual components and are thus searched. The parts b) and c) of claim 41 have however, not been searched, since they respectively refer to mammalian homologues/orthologues functionally active fragments or derivatives of the proteins (a) and (b) with furthermore undefined single or multiple substitutions deletions and/or additions. These latter have no support or teaching (Art. 84 and 83 EPC) and have therefore not been searched for similar reasons to those given above. The same applies to any dependent claimed subject-matter (claims 41-45).

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.